Tenth Annual Meeting of the Society for Text and Discourse

Lyon, France
July 19 – 21, 2000

History Learning as Literacy: Cognitive Text Research and Instructional Applications. C. A. Perfetti

LSA: A Psychological Theory of Meaning and its Applications. W. Kintsch

Modeling Online Construction of a Multidimensional Situation Model in the Landscape Model of Comprehension. Y. Tzeng, P. van den Broek & R. Zwaan

Situation Model and Causal Contradictions: A Distance Effect for Narrative Characters' Properties. N. Campion & D. Martins

Updating of Spatial Changes in Situation Models. M. Rinck, K. Wolf & J. Hasebrook

The Influence of Focus on Updating a Mental Representation. H. van Oostendorp & C. van der Puil

Updating and the Reactivation of Situation-Model Information. R. A. Zwaan, C. J. Madden & R. A. Stanfield

The Effectiveness of Tutorial Dialog in an Automated Conversational Tutor. K. Link, V. Pomeroy, R. DiPaolo, S. Rajan, B. Klettke, L. Bautista, R. Kreuz, A. Graesser & The Tutoring Research Group

The Dialog Advancer Network: A Mechanism for Improving AutoTutor's Conversational Skills. N. Person, A. C. Graesser, D. Harter & The Tutoring Research Group

Repetition in Discourse: A Linguistic Strategy Signifying Involvement in Children's Conversational Dialogues with a Literary Work. R. Horowitz & C. Cummings

Accessability, Duration, and Modeling the Listener in Dialogue. E. G. Bard & M. P. Aylett

Achieving Understanding in Multiparty Interactions. A. H. Anderson, J. Mullin, E. Katsavras, R. McEwan, E. Grattan & P. Brundell

Psychology, Discourse and Ideology. T. A. van Dijk

The Acquisition of Information Search Skills in 9 to 13 Year-Old Students. J-F. Rouet & C. Chollet

LSA in the Classroom: Automatic Feedback for Learning Summarizing Skills. E. Kintsch & D. Steinhart

The Use of Narrative in Argumentation. J. F. Voss & J. van Dyke

The Understanding Of Nominal Metaphor. C. Tijus, B. Pudelko, E. Hamilton, & D. Legros

The Costs and Benefits of Metaphor. I. A. Noveck & M. Bianco

Prosodic Correlates of Text Structure. H. J. N. den Ouden, L. G. M. Noordman & J. M. B. Terken

Analysis of Descriptive and Narrative Texts Written by French and Dutch Children. L. Chanquoy & J. Schilperoord

Sources of Reading Comprehension Difficulties in Children: On-line Processing of Cohesion in Listening Comprehension. M-F. Ehrlich & H. Megherbi

Transitory Situation Model and Linear Procedural Model: Their Roles in Comprehension of Word Problems Statments at Primary School. D. Coquin-Viennot & S. Moreau

Metacognitive Knowledge in Text Comprehension: Some Issues in Development and Individual Differences. P-E. Eme & C. Haro

The Development of Analogical Processes in Learning to Read in French. I. Brun, J. Faure, J. Ecalle & A. Magnan

Development of Syntactic Connexion in French Children's Narrative and Expositive Texts. A. Viguié

Phonological Recoding in Deaf Children. L. Paire-Ficout, S. Colin & A. Magnan

The Influence of the Phonological and Orthographic Characteristics of French During Reading Acquisition.J. B. Aimar & A. Magnan

Select-a-Kibitzer: A Multi-Agent Architecture for Giving Feedback on Student Compositions. P. Wiemer-Hastings & A. C. Graesser

ETAT: A Tool to Analyse Expository Text Coherence. E. Vidal-Abarca, H. Reyes, R. Gilabert, J. Calpe, E. Soria & A. C. Graesser

Insights into the Interactive Process of a Computerized Tutor:A Focus Group Study. R. DiPaolo & S. Rajan

Short Responses in Human and Intelligent Tutoring Systems S. Rajan, D. Harter & A. C. Graesser

Verb-based Classification of Abstract Concepts.K. Wiemer-Hastings & A. C. Graesser

AutoCoder: An Intelligent Assistant for Coding Protocol Data. J. Earwood, M. A. Durbin & R. M. Golden

Conceptual Change via Learning from a Text. A Connectionist Model. V. Sanjose, O. M. Padilla & E. Vidal-Abarca

The Effects of Information Presentation Style on Question Generation.S. D. Craig, B. Gholson, M. Ventura & The Tutoring Research Group

Advertising Discourse: A Critical Approach to Chinese Advertisements. V. T-H. Chang

Language-Related Causes of Communication Breakdown in Medical Practice. M. Ibba

More than a Stately Dance: Dialogue as a Reaction Time Experiment. E.G. Bard, M.P. Aylett & M. Bull

Literacy is More than Learning to Read: Information Retention from TV News. S. Gulgoz, F. Goksen & C. Kagitcibasi

Functions of Honorifics and Topic Manipulation in "Discourse Politeness" as Unmarked Politeness: From the Analyses of Japanese Conversations. M. Usami

The Chechnya War: Soviet Political Discourse Revisited. T. Tsyrendorjieva

Understanding the Ironic Language. C. Munch

Are you Trying to be Funny? The Importance of Exaggeration on Discourse Goal Clarity. C. O. Stewart & R. J. Kreuz

Interactions on the Web Between Reviewers and Authors: Towards Reviewing an Article for the e-Journal "JIME". A. Chevalier, N. Bonnardel & A. Piolat

The Impact of Verbal Information on the Aesthetic Experiences to Visual Art. K. Millis, K. Schaefer & B. Salmon

Conjoined Effects of Long-Term Working Memory Representation, Text Difficulty and Reading Purpose on Reading Strategies in Text Problem Comprehension. M. A. Schelstraete & I. Walschaerts

The Effect of Verbal Context on Picture Recognition: Initial Support for Perceptual Symbol Theory. R. A. Stanfield & R. A. Zwaan

Long-Term Working Memory: Some Empirical Results. C. Bellissens & G. Denhière

Perspective Effects on Fixation Times and Memory for Text. J. Kaakinen, J. Hyönä & J. M. Keenan

The Effects of Testing and Delay on Recall and Learning of Information from Texts. S. Gulgoz, M. E. Aktunc, D. Senol & H. A. Van

The Study of Locative Sentences from Corpora of Narratives. M. de Vega, M. J. Rodrigo, M. Ato, D. Dehn & B. Barquero

Processing Figurative Language in the Underspecification Model. S. Frisson & M. Pickering

Metaphor and the Space Structuring Model: Evidence from Event-Related Brain Potentials. S. Coulson

Figurative Language: Retention vs. Suppression. R. Giora

Convention, Form, and Figurative Language Processing. B. F. Bowdle & D. Gentner

Cognitive Evidence for a Parameterization of Cohesion and Coherence relations. M. Louwerse

Coherence and Evidence in Testimony Evaluation on Incest Narratives. B. Klettke & A. C. Graesser

Generating Inferences from Scientific Text. J. Wiley & J. L. Myers

Idiom Processing : Effect of Compositionality. S. Caillies , K. Butcher

Why Should We Use Connectives in Discourse? Overspecification of Referential Expressions in Instructive Discourse. A. Arts, A. Maes, L. G. M. Noordman & C. Jansen

Causal Factors that Influence Story Agents, Blame and Responsibility B.A. Olde & A. C. Graesser

Measuring Causal Force in Causal Chains. A. Guha & J-P. Rossi

Comparing the Contribution of the Visuo-Spatial Sketchpad and Phonological Loop in the Written Recall of a Story. T. Olive, V. Li Calzi & A. Piolat

Diagnosis of Cognitive Functioning During Text Reading and Comprehension by Adolescents. H. Thomas

Representation and Processing of Contrastive Information in discourse Comprehension. J. M. Lee, S. Choi, J. G. Lee, & K. H. Cho

Inter-Individual Differences in the Integration Processes of Text Comprehension: Do they Influence the Use of the Leading Edge Strategy and Comprehension? L. Demanet, M. A. Schelstraete & M. Hupet

Representation of Simple Arguments During Reading. A. Britt, J. F. Rouet & C. A. Perfetti K/p>

Semantic Macrostructure Effect in Accessing Relevant Prior Information during Text Reading. S. Montoya, G. Denhière & T. Baccino

Which Role the Position of Lexical Explanations Plays in Text Comprehension?. E. Jamet & O. Le Bohec

Textuality and the Rhyming Principle in Narrative Superstructure. H. Masuda

Building a Mental Model from Text and Pictures: The Role of the Visuo-Spatial Working Memory in the Memory of Pictures. V. Gyselinck & M. F Ehrlich

Story Comprehension in Patients with Damage to Prefrontal Cortex. T. Zalla/, M. Phipps, M. Colvin & J. Grafman

Functional Hemispheric Asymmetry in Voicing Feature Processing in Reading. N. Bedoin & H. Chavand

"My Recaller is on Vacation" Can memory books change discourse patterns in residents with dementia? K. Dijkstra, M. Bourgeois, G. Petrie, L. Burgio & R. Allen-Burge

The Effect of Context and Categorizability on Verbal Learning T. Guthke, E.C. Ferstl & A. Hauptmann

Non-aphasic Language Deficits After Brain Injury: A Comparison of Comprehension Skills of Patients with Different Discourse Production Profiles E. C. Ferstl, T. Guthke, F. Siebörger & D. Y. von Cramon

---

History Learning as Literacy: Cognitive Text Research and Instructional Applications

C. A. Perfetti

University of Pittsburgh (U.S.A.)

perfetti+@pitt.edu

Amidst much published confusion about the nature of literacy, a simple distinction between basic literacy and extended literacy is useful. Basic literacy is fundamentally about the acquisition of knowledge of written words, their forms and meanings, and secondarily about the application of language abilities and general knowledge to the printed word. Extended literacy goes beyond these basics to engage students in literacy–based problem solving, writing and argumentation. In considering "the literacy problem", it is important not to confuse the one with the other. Basic literacy is a prerequisite for extended literacy and different obstacles to competence arise at the two levels. Elsewhere (e.g. Perfetti, 1998), I have focused on the lexical foundations of basic literacy. Here the focus is extended literacy—demanding text environments that require reasoning and problem solving. What processes and skills are required by multiple-text problem-oriented reading? What kinds of learning environments can support the acquisition of these skills?

The study of history offers a distinctive, rich literacy environment. It centers on documents-letters, treaties, notes, official records, diaries--as well as textbooks. Thus, instruction that makes good use of this rich text environment has the potential to support broad-based literacy skills that may extend beyond history classrooms to other cases of text learning, reasoning, and writing. Indeed, in the United States, the National Standards for United States History include thinking skills that "enable students to evaluate evidence, develop comparative and causal analyses, interpret the historical record, and construct sound historical arguments and perspectives on which informed decisions in contemporary life can be based" (National Standards for United States History, 1996, p. 2). The problem is that history instruction in the United States may not do enough to promote these examples of extended literacy: "The typical history classroom is one in which they listen to the teacher explain the day's lesson, use the textbook, and take tests. Occasionally they watch a movie. Sometimes they memorize information or read stories about events and people. They seldom work with other students, use original documents, write term papers, or discuss the significance of what they are studying" (p. 194). This dramatic gap between the literacy potential of history study and the bleak picture of the actual practice of history study (at least in the U.S.) suggests an interesting problem for the application of cognitive science. We–my collaborators, J.F. Rouet, Peter Foltz, Julie Van Dyke, Gareth Gabrys, and especially M. Anne Britt and I–have focused on history learning as a case of extended literacy. In addition to the interesting theoretical issues for text learning that history affords, we hoped to address the gap between history's potential and practice mentioned above. Our project includes theoretical, empirical, and instructional components, as summarized below: The Documents Model is a framework for describing how readers represent multiple documents that deal with a common problem space (Perfetti, Rouet & Britt, 1999). The framework includes a Documents Space where documents are represented and interlinked by Rhetorical Predicates to form the learner's Intertext Model. They also are linked to a Situations Model, representing the learner's understanding of one or more temporal-causal scenarios that are the "story" (or stories) revealed in the texts. Thus, the theoretical formulation represents both causal networks (Trabasso & van den Broek, 1985) and strict text representations,

distinct from inferred situations (van Dijk & Kintsch, 1983). It allows a framework for student understanding. For example, many students achieve a well developed situation model with only weak links to one or more texts. A higher standard is seen in students who achieve an articulated intertext model as well as a situation model. Most sophisticated is the student who can entertain two or more situation models linked to specific texts. One assumption of the theory is that learning from multiple documents sharpens the distinction between text and situation for the reader. This implies that multiple documents can promote a more flexible representation of the situation, and, indeed students who write essays based on multiple texts transform the text information more than students with access to a single text (Wiley & Voss, in press).

Empirical component

Our studies have examined sustained learning over time with real texts, experimental manipulations of document characteristics, and other variables. They suggest, among other things, that both younger and college age readers readily acquire basic temporal-causal structures that comprise the narrative heart of many historical texts. And college age readers can construct flexible representations that reflect the integration of information across texts with contrasting perspectives (Perfetti, Britt, & Georgi, 1995). But students at high school and even college levels lack clear perceptions on the uses of documents and the evidential privileges they entail for arguments. We have observed a preference among regular-class American history students for textbooks, not only in use but in trustworthiness. We have observed differences among students, including discipline-background differences (Rouet, Favart, Britt, & Perfetti, 1998) in their appreciation of document privilege. Other studies have examined aspects of the Documents Theory, notably the trade-off between integration of information across documents and the separation of source information.

A Computer Document Tool

In light of our studies of students learning history in multiple text environments, we concluded that high school students could benefit from help in the use of multiple documents. We developed the Sourcer's Apprentice (SA) to assist students in evidence seeking and evaluation skills and to foster an awareness of document type and document privilege (Britt, Perfetti, van Dyke & Gabrys, in press). The SA incorporates instructional design principles and reflects research on obstacles to high school students' use of documents (Wineburg, 1991). For example, it engages the student directly in text-based problem solving, one of the goals of extended literacy. It displays a bookshelf with documents from which the student can select and take notes while addressing a problem. Students can use the tool as a tutor for direct instruction or as a problem-solving writing environment. Several studies on the effectiveness of the SA as a learning tool have produced positive results in usability and effectiveness. A recent study has suggested more causally coherent essays are written by high school students who have used the SA.

Conclusion

Have we learned something about how to promote a wider realization of the potential of history learning for extending literacy? In a project with many facets, there are many small conclusions: About the privilege of temporal-causal representations as mental models for students; about the obstacles and opportunities this preference for narrativity present to real document study; about the theoretical question of integrating content while separating sources; about the ability of a simple tool, the SA, to support student learning of document use; etc. At a more general level, we

believe that history as extended literacy is attainable in a wide range of students who have achieved basic literacy. The special demands on problem-oriented multiple text learning, however, may require some direct and explicit instruction for many students. The Sourcer's Apprentice is one example of how such instruction might occur.

LSA: A Psychological Theory of Meaning and its Applications

W. Kintsch

University of Colorado at Boulder (U.S.A.)

wkintsch@psych.colorado.edu

LSA is a mathematical technique that generates a high-dimensional semantic space from the analysis of a large corpus of written text. The technique was originally developed in the context of information retrieval (Deerwester, Dumais, Furnas, Landauer, & Harshman, 1990) and was adapted for psycholinguistic analyses by Landauer and his colleagues (Landauer & Dumais, 1997; Landauer, Foltz, & Laham, 1998; Landauer, 1999).

LSA must be trained with a large corpus of written text. The raw data LSA are meaningful passages and the set of words each contains. A matrix is constructed whose columns are words and whose rows are documents. The cells of the matrix are the frequencies with which each word occurred in each document. The data upon which the analyses reported below are based consist of a training corpus of about 11 million words (what a typical American school child would read from grade 3 through grade 14), yielding a co-occurrence matrix of more than 92,000 word types and more than 37,000 documents. Note that LSA considers only patterns of word usage; word order, syntax, or rhetorical structure are not taken into account.

Word usage patterns, however, are only the input to LSA which transforms these statistics into something new - a high-dimensional semantic space. LSA does this through dimension reduction. Much of the information in the original pattern of word usage is accidental and inessential. Why did an author choose a particular word in a specific place rather than some other alternative? Why was this particular document included in the corpus rather than some other one? LSA discards all of this excess information and focuses only upon the essential semantic information in the corpus. To tell what is essential and what is distracting information, LSA uses a standard mathematical technique called singular value decomposition, which allows it to select the most important dimensions underlying the original co-occurrence matrix, discarding the rest. The matrix is decomposed into components associated with its singular values, which are ordered according to their importance. The 300 most important components define the semantic space. The dimensionality of the space is chosen empirically: A (roughly) 300-dimensional space usually compares best with human performance.

LSA thus makes the strong psychological claim that word meanings can be represented as vectors in a semantic space of approximately 300 dimensions. But not only word meanings are represented as vectors in this space, documents are similarly represented as well. And new documents - sentences, paragraphs, essays, whole book chapters - can also be represented as vectors in this same space. This is what makes LSA so useful. It allows us to compare arbitrary word and sentence meanings, determine how related or unrelated they are, and what other words or sentences or documents are close to them in the semantic space. A word of caution is necessary here: LSA knows only what it has been taught. If words are used that did not appear in the training corpus, or which are used differently than in the training corpus, LSA, not unlike a person, does not recognize them correctly or at all.

The measure that is used to calculate semantic relatedness is the cosine between two vectors. As a first approximation, readers unfamiliar with this concept may think of cosines as analogous to correlation coefficients. The cosine varies from -1 to +1, +1 denoting identity and 0 denoting unrelatedness. Most cosines between words are positive, though small negative values are common (the average cosine for randomly chosen word pairs is .02, with a standard deviation of

.06). The more closely two words are related semantically, the higher their cosine. For instance, the singular and plural forms of a sample of 100 common nouns had a mean cosine of .66, with a standard deviation of .15.

A second measure that is often useful is the length of a vector, which, like the cosine, is defined mathematically. Intuitively, the vector length tells us how much information LSA has about this vector. Thus, the length of sentence vectors is generally greater than the length of word vectors,

and the length of paragraph vectors is even greater. Words that LSA knows a lot about (because they appear frequently in the training corpus, in many different contexts) have greater vector lengths than words LSA does not know well. Thus, horse has a vector length of 2.49, while porch has a vector length of .59. Function words that are used frequently in many different contexts have low vector lengths (the and of have vector lengths of .03 and .06, respectively, and their cosine is .99 - LSA knows nothing about them and cannot tell them apart since they appear in all contexts).

All we can do, however, is compare one vector with another. Inspecting the 300 numbers that compose it tells us little, for the dimensions of the semantic space are not identifiable. The only way we can tell what a given vector means is to find out what other words or sentence vectors are close to it. Thus, we can ask LSA to list the words closest to a given vector in the semantic space. The semantic neighborhood of a word tells us a great deal about the word. Indeed, we shall make considerable use of semantic neighborhoods below.

Often we have some specific expectations about how a vector should be related to particular words or phrases. In such cases it is most informative to compute the cosine between the vector in question and the semantic landmark we have in mind. In most of the examples discussed below when we need to determine what a vector that has been computed really means, it will be compared to such landmarks. Suppose we compute the vectors for horse and porch. To test whether what has been computed is sensible or not, we might compare these vectors to landmarks for which we have clear-cut expectations. For instance, the word gallop should have higher cosine with horse than with porch (the cosines in fact are .75 and .10, respectively), but the word house should have a higher cosine with porch than with horse (the cosines are .08 for horse and .65 for porch). This is not a very powerful test, but it is intuitively compelling and simple. What the particular landmarks are is not terribly important, as long as we have clear shared semantic expectations. Someone else might have chosen race instead of gallop, or door instead of house, or many other similar word pairs, with qualitatively equivalent results.

Readers can make their own computations, or check the ones reported here, by using the web site of the Colorado LSA Research group: http://lsa.colorado.edu. First select the appropriate semantic space and dimensionality. The semantic space used here is the "General Reading through First Year of College" space with 300 dimensions and term-to-term comparisons. To find the semantic neighborhood of horse, one types "horse" into the Nearest-Neighbor-box and chooses "pseudodoc". To find the cosine between horse and gallop, one types "horse" and into one box and "gallop" into the other box of the One-to-Many-Comparison.

LSA has proved to be a powerful tool for the simulation of psycholinguistic phenomena as well as in a number of applications that depend on an effective representation of verbal meaning. Among the former are Landauer and Dumais (1997), who have discussed vocabulary acquisition as the construction of a semantic space, modeled by LSA; Laham's (1997) investigation of the emergence of natural categories from the LSA space; and Foltz, Kintsch and Landauer's (1998) work on textual coherence. To mention just three of the practical applications, there is first, the use of LSA to select instructional texts that are appropriate to a student's level of background knowledge (Wolfe, Schreiner, Rehder, Laham, Foltz, Landauer, & Kintsch, 1998). Second, LSA has been used to provide feedback about their writing to 6th-grade students summarizing science or social science texts (E. Kintsch, Steinhart, Stahl, Matthews, Lamb, & the LSA Research Group, in press). The application of LSA that has aroused the greatest interest is the use of LSA for essay grading. LSA grades the content of certain types of essays as well and as reliably as human professionals

(Landauer, Laham, Rehder, & Schreiner, 1997). The human-like performance of LSA in these areas strongly suggests that the way meaning is represented in LSA is closely related to the way humans operate. The present paper describes an LSA-based computational model, which accounts for another aspect of language use, namely, how meaning can be modified contextually in predication. The model is discussed first and illustrated with some simple examples of predication. Then the model is used to simulate several more complex kinds of language processing.

Modeling Online Construction of a Multidimensional Situation Model in the Landscape Model of Comprehension

Y. Tzeng*, P. van den Broek & R. Zwaan***

*National Chung Cheng University (Taiwan)

**University of Minnesota (U.S.A.)

***Florida State University (U.S.A.)

ttcytt@ccunix.ccu.edu.tw

Readers gain a true understanding of a text by building a situation model. The study of situation models thus is central to understand comprehension processes (see the seminal work by van Dijk & Kintsch, 1983). A new development in the study of situation models is the proposal of the Event-Indexing model (Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995; Zwaan & Radvansky, 1998). This model hypothesizes that readers build situation models by monitoring at least five event dimensions: Protagonist, time, space, causation, and intentionality. By integrating information from these dimensions, readers gradually update their mental representations and build a richly interconnected coherent network or situation model.

Experimental data generally have supported the predictions of the event-indexing model (e.g., Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995; Zwaan & Radvansky, 1998; Zwaan, Radvansky, Hilliard, & Curiel, 1998). However, converging evidence is necessary to test different aspects of this model. One important feature of a good scientific theory is the display of internal consistency. A rigorous way of validating a model's internal consistency is to examine whether it can be implemented in a computational architecture and still can reasonably account for human performance. Therefore, one purpose of the current study was to examine whether the event-indexing model is detailed and consistent enough to be implemented computationally. A second purpose of this study was to test whether a particular computational model of comprehension, the Landscape model, is able to model online processes of situation models.

The Landscape model is a connectionist model designed to capture the relations between online processes and offline representation of comprehension (Tzeng, 1999; van den Broek, Young, Tzeng, & Linderholm, 1999; van den Broek, Risden, Fletcher, & Thurlow, 1996). The model assumes that readers routinely make causal and referential inferences during the processes of reading and build a coherent representation through a delta learning rule. Furthermore, the Landscape model has a recurrent property, cohort activation, which allows the model to dynamically incorporates readers' activation state from the previous reading cycle into the current mental representation. A cohort is a group of reading elements that are related because they were co-activated at some time during reading. An important property of cohort is that if any part of a cohort is activated the rest of the cohort will be activated as well, to a less degree. The result of reading comprehension by the Landscape model is a representation that emerges from the landscape of fluctuating activations. The landscape itself reflects from various sources of activation. The model has been found to account for a high proportion of readers' recall and to do so across a wide range of texts (Tzeng, 1999; van den Broek, et. al., 1996). Moreover, the model shows potential improvements over the Construction-Integration model (Kintsch, 1988; see Tzeng, 1999).

Accounting for offline recall data only explains the products of reading. The question remains whether the model can predict online comprehension processes. Thus, modeling online reading time data predicted by a theory of situation model becomes a robust challenge for the Landscape

model. The event-indexing model was implemented within the Landscape model architecture, with a focus on the idea of continuity between sentences. According to the event-indexing model, sentences vary in their degree of continuity, as determined by the number of dimensions in which consecutive sentences overlap. Specifically, sentence continuities vary along the five dimensions of situation models. The continuity of sentences among these five dimensions will affect readers' processing speed and, hence, sentence reading times. The event-indexing model predicts that sentences with stronger continuity are easier to process than those with less continuity. As a result, a sentence sharing more dimensions with its previous sentence will have less reading time than another sentence sharing fewer dimensions with its previous one.

Sentence reading time data reported in the second experiment of Zwaan, Radvansky, Hilliard, and Curiel (1998) were adopted for modeling purpose. These reading times were adjusted for the number of syllables. The number of overlap along the five dimensions of situation models was determined for consecutive sentences and used as an indicator of continuity. A sentence can overlap with another by zero to five dimensions therefore the range of continuity was between 0 to 5. These continuity measures were transformed into an input matrix and submitted to the Landscape model. The amount of new activation was computed for each reading cycle by subtracting the total amount of activation of the earlier cycle from the total amount of activation of the current cycle. These activation differences, representing the continuity between sentences, were correlated with sentence reading times by human readers. The resulting correlation coefficients indicate to what extent the Landscape model can model these online reading times.

The observed correlation strengths between activation differences (sentence continuity) and reading times ranged from .46 to .75 with a mean of .57 (all ps < .05). This pattern of results suggests that the Landscape model aptly simulate the online processes predicted by a theory of situation models. It also indicates that the event-indexing model is well specified and it is logically consistent to be captured by a computational model. Thus, this study simultaneously has provided evidence for both the Landscape and the event-indexing theories.

Situation Model and Causal Contradictions:

A Distance Effect for Narrative Characters' Properties

N. Campion* & D. Martins**

*University of Paris XI (France)

**University of Paris X (France)

Campion@limsi.fr

We addressed how readers will detect and process a causal contradiction between the properties of a narrative character and the actions performed by that character. Albrecht and O'Brien (1993) already showed that the contradiction was detected when three sentences intervened between the description of the character's properties and the description of its actions. Thus, Albrecht and O'Brien concluded that readers construct and update a situation model that maps the causal properties of the described situation. Here we looked at what would happen if the contradictory actions were stated immediately after the character's description.

Method

Material. Eighteen experimental texts were written. After an introduction of two sentences, an elaboration part of the text was made of three sentences which expressed specific properties of a main character. These elaboration parts were followed by two sentences which expressed the critical actions performed by the main character. Three coherence conditions were created by varying the character's properties, although the critical actions remained the same. In the Coherence and Incoherence conditions, our knowledge allowed or not to establish a causal relation between the character's properties and the critical actions. In the Neutral condition, there were no causal relation between the properties and the critical actions. Distant and close conditions were also created by adding or not three filler sentences between the elaboration part and the critical actions. In the distant condition, the character's properties were considered as backgrounded by the reading of the three filler sentences and no more present in readers' working memory.

Procedure. Seventy-twelve students read the experimental texts which were randomly mixed with filler texts. The reading was made sentence by sentence on a computer screen and at a free speed. In the experimental design, a coherence factor (Coherent, Incoherent, Neutral) varied within subjects and a distance factor (Short, Long) varied between subjects.

Results

The reading times for the two critical action sentences are presented in Table 1. Globally, the reading times were significantly shorter for the first sentences than for the second sentences (at the critical threshold of $p < .05$ as for the following results). Separate analyses of variance (ANOVA) were performed for each of these sentences and showed exactly the same pattern. For the first or second sentences, there was a significant interaction between the factors Coherence and Distance. Planned comparisons confirmed that the factor Coherence had no significant effect in the Close condition, but had a significant effect in the Distant condition. In the Distant condition,

the reading times were significantly shorter in the Coherent condition than in the Incoherent or Neutral conditions. Our results first confirmed that readers detect the causal discrepancy between the properties previously attributed to a narrative character and an action performed three sentences later by that character. This is attested by readers difficulties to process the contradictory actions and corroborate their use of a situation model to update characters properties. However, our results also showed that when the contradictory action was stated immediately after the character's description, readers had no more apparent difficulties to process the contradictory action. Thus it seems that a situation model first need to be backgrounded by the process of intervening material for a causal contradiction being perceived as such. A plausible explanation is that, when the situation is first introduced in a text and when its model is still activated in reader's working memory, the contradictory actions are interpreted as an interesting paradox (Schank, 1979) or as an initial outcome introduced by the author. Therefore the readers would notice the contradiction, but would adopt a "wait and see strategy", considering that further explanation will be stated in the following of the text. In the distant condition, however, the contradiction was perceived but was now considered as unacceptable. In such case, readers probably considered that a conventional rule of consistency would be violated by the introduction of the contradictory actions at that moment without previous explanation. Therefore readers try to solve themselves the contradiction. They may check their text base representation to find information that they might have neglected while building their situation model or more fully activate their situation model (O'Brien, Rizzella, Albrecht & Hallebran, 1998). They could also try to infer an elaboration that solves the contradiction. Finally, when a situation model has been backgrounded in a fixed state, because of filler information, that state would become a reference point for the reader and thus, without a previous explanation, the content of the model could be updated by coherent completion but not by radical change. On the contrary, the situation model could still be easily modified when it is initially activated in working memory and has not yet been backgrounded. It should however be noticed that this last conclusion is probably limited to characters properties that are described as a usual state. It has been showed to be invalid for unfocused spatial relation (O'Brien & Albrecht, 1992) and for stable knowledge about the world that are incongruent with a character goal ("Swim and sunbathe are practiced in Florida and not in Alaska", Huitema, Dopkins, Klin & Myers, 1993).

Table 1. Mean reading times (ms) for the two critical action sentences in the Coherence and Distance conditions

		Coherent	Incoherent	Neutral
	First	3426	3500	3555
Short	Second	3043	3154	3196
Long	First	3355	3815	3548
	Second	3159	3454	3267

An example of text: Francoise was living close to many little shops. She preferred to do her shopping in those shops than in the supermarket.

Coherent: Francoise was the best client in a little butchery because she was found of red meat. She specially liked beef and here favorite steaks were the ribs, the sirloin and the roast. She always eat her meat very little cooked and was telling everybody to eat meat underdone.

Incoherent: Since many years Francoise had stopped to go to the butcher because she became a strict vegetarian. She would never eat any beef whatever the pieces. Only the idea of eating ribs, sirloins or roast beef was disgusting her very deeply.

Neutral: Francoise was probably the best client in a little bakery because she was found of their special bread. She was specially found of their bread made with three kinds of cereal: wheat, rye and bran. She always eat her bread as soon as possible when it was still warm.

Filler: Francoise was not living so far from downtown and needed to make a little bit of sport. Evry weekend, Francoise appreciated to walk down town when it was not too cold. It was taking exactly an hour to her to go from her door to the most central building of the town.

Critical 1: That night, Francoise joined a friend in a restaurant and ordered for herself an enormous beef steak.

Critical 2: She eat her meat with pleasure and found it particularly tasteful and juicy.

End: After the diner Francoise proposed a travel plan to her friend. She wanted to cross Africa by foot with her friend. The friend of Francoise couldn't imagine that Francoise was serious.

Updating of Spatial Changes in Situation Models

M. Rinck*, K. Wolf* & J. Hasebrook**

*Technical University of Dresden (Germany)

**Bankakademie Frankfurt (Germany)

rinck@Rcs1.urz.tu-dresden.de

Current theories of text comprehension assume that readers build situation models of the information described by texts. Situation models constitute the level of text representation associated with "deep" processing, and serve to integrate the information stated in a text with general information supplied by the readers' world knowledge. Many experiments investigated how readers update spatial aspects of their situation model during narrative comprehension, following a paradigm introduced by Morrow and his colleagues (e.g., Morrow, Greenspan, & Bower, 1989). In this paradigm, participants first memorize the spatial layout of a building including its rooms and a number of critical objects, and later read stories about characters moving through the building in pursuit of some goal. These experiments have shown that readers focus their attention on the protagonist of the story. A primary consequence of the readers' focus of attention is that known objects close within the model to the current focus become more activated in memory, so that readers can readily refer to them or retrieveinformation about them to answer questions. The term spatial distance effect was coined for this result: The closer an object is to the protagonist of the story, the more accessible it is in memory.

Despite many replications of the distance effect (e.g., Rinck & Bower, 1995; Rinck, Hdhnel, Bower, & Glowalla, 1997), the type of spatial updating investigated in these studies is actually very limited. Except for the location of the protagonist, all spatial aspects of the situation model remain constant. For instance, neither the rooms of the memorized building nor the objects in them change over the course of the experiment. Therefore, one goal of the two experiments described here was to find out if readers also update more drastic changes of the spatial scenario, namely reconstructions of the previously memorized building. These reconstructions involved the addition and deletion of walls separating rooms of the building. The second goal was to find out if these changes affect the spatial distance effect observed in previous studies. Presumably, the distance effect should be stronger if the protagonist and the probed object are located in different rooms, that is, if a separating wall exists between them (see Rinck et al., 1997). Third, the present experiments aimed at answering these questions with more naturalistic materials than had been previously used. Therefore, a naturalistic long text, similar to a soap opera involving a cast of main characters was read by the participants, and the layout learning procedure was part of reading the text.

Both experiments followed the same basic procedure while participants were reading the text sentence by sentence in a self-paced manner on the computer screen. In the first part of the text, the main characters and the spatial setting were introduced: A group of researchers who had to move into a building which had to be remodeled before it could be used as a research building. In this part, the original remodeling plan was described using both verbal and map information. In the second part of the narrative, participants read that two critical walls had been forgotten during remodeling. Again, this was described by both verbal statements and a map showing all existing and missing walls. After this, a set of episodes described how the protagonists worked and lived in the incomplete building. In the third part, the text described that the building had to be remodeled again to introduce the previously forgotten walls. However, the result of this second remodeling attempt was unsatisfactory again, yielding the four different experimental conditions of "wall status": One critical wall was correct in that it had been put in during the first remodeling and was

still there after the second (always existent wall), one wall had been put in correctly first but was accidentally removed during the second remodeling (removed wall), one wall had been missing after the first remodeling and was added correctly during the second (added wall), and one wall was missing both after the first and the second remodeling (always missing wall). In the final part of the text, participants read another set of episodes describing events and actions of the protagonists in the still unsatisfactory building.

As part of these final episodes, three types of experimental sentences were used to explore readers' updating of the spatial changes involved in the two remodeling events. First, "watch sentences" described how the protagonist saw an event in the adjacent room, thereby following the inconsistency paradigm introduced by O'Brien and Albrecht (1992). These sentences made sense only when the wall between the protagonist's location room and the adjacent room was missing. In both experiments, reading times of these sentences were longer if they were inconsistent with the situation model, that is, if the wall existed. If the wall had been removed or had been missing all the time, reading times were significantly shorter, indicating that readers incorporated the existence of the walls into their situation model. Second, "motion sentences" stated that the protagonist walked from his current location room into the adjacent room. Earlier studies have shown that reading times of sentences such as these are correlated with the length of the protagonist's path in the situation model. Indeed, in both experiments described here, reading times were significantly shorter if there was no wall between the location room and the adjacent room, reflecting the fact that the protagonist could walk directly into the adjacent room rather than having to take a route via the hallway. Thus, reading times of these sentences indicate that readers used their knowledge about the existence of walls to represent the protagonist's movements. Third, "anaphoric target sentences" stated that the protagonist thought about or remembered an object located in the location room or in the adjacent room. Given the previously found spatial distance effect, reading times of these sentences should be shorter if they refer to the spatially close objects in the location room compared to the more distant objects in the adjacent room, reflecting the fact that the former are more accessible in memory than the latter. However, no such difference was found in Experiment 1. We reasoned that this might have occurred because - unlike previous experiments - the text did not refer repeatedly to the protagonist's location. Therefore, the text was slightly altered in Experiment 2 to mention the protagonist's location more often. This yielded a weak, but significant distance effect. Thus, it seems that the spatial distance effect might not be so very general after all, and that readers only pay close attention to the protagonist's spatial location if the text or the instructions make it particularly salient.

References

Morrow, D. G., Greenspan, S. C. & Bower, G H. (1987). Accessibility and situation models in narrative comprehension. Journal of Memory and Language, 26, 165-187.

O'Brien, E. J., & Albrecht, J. E. (1992). Comprehension strategies in the development of a mental model. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 777-784.

Rinck, M. & Bower, G. H. (1995). Anaphora resolution and the focus of attention in situation models. Journal of Memory and Language, 34, 110-131.

Rinck, M., Hdhnel, A., Bower, G. & H., Glowalla, U. (1997). The metrics of spatial situation models. Journal of Experimental Psychology: Learning, Memory, and Cognition, 23, 622-637.

The Influence of Focus on Updating a Mental Representation

H. van Oostendorp & C. van der Puil

Utrecht University (The Netherlands)

H.vanOostendorp@FSS.UU.NL

An important process in text comprehension is the updating of mental representations during reading, that is, the process of transforming knowledge already represented in a readers' memory in the light of new, correcting information (van Oostendorp, 1996; Johnson & Seifert, 1999; van Oostendorp & Bonebakker, 1999). For instance, van Oostendorp (1996) had people read a news report about a military intervention in Somalia followed by a second related text. This second text contained transformations of facts mentioned in the first text. Performance on a later inference task showed that little updating occurred. Furthermore, it appeared that changes that were more important to the situation described were less updated. It was suggested that central parts of a situation model may be less easily updated than peripheral parts (cf. Chinn & Brewer, 1993). In this paper we studied whether the central part of a mental representation is less updated when this information is in focus. Focus was manipulated by letting readers compare one text with another along some dimension. The way we examined whether the representation has been updated is based on a method, recently introduced by Ferstl and Kintsch (1999), the Cued Association task. With this task subjects are presented with a word (cue) and they are asked to provide an association to it. The number of times a keyword is given as a response to another keyword is used as a measure of relatedness between these two words. Based on these data an asymmetric proximity matrix is calculated for each subject, which is formally equivalent to a network structure. The keywords represent the nodes and the strengths of links are determined by the response order. In these network structures it is possible to discern concepts that are highly related to each other, and concepts less strongly related. Useful notions for measures of centrality are here the in-degrees and out-degrees of concepts (the summation of all the links going into a node or out of a node, respectively; see Ferstl & Kintsch, 1999). The central part of the mental representation was here defined by the area containing concepts with the highest inner cohesion (that is, having the highest number of links between selected concepts). Subsequently, we examined the similarity of the central parts of the networks of the subjects in the focus group before and after reading the texts and compared that with the similarity in the non-focus group. In line with the preceding study (van Oostendorp, 1996) we hypothesized that the central part of the mental representation is less updated when that information is in focus compared to the updating of this information when it is not in focus. We predicted, thus, that the similarity of the central part of the network structures will be higher in the first case (when it is in focus) than in the latter case (when it is not in focus).

Method

Subjects (40 university students) received two related texts about the Spanish terrorist movement ETA. The second text contained a number of important changes or corrections of facts compared to the first ETA text. Each text could be a typical newspaper article (about 500 words each). Both texts contained two kinds of information. First, information about attacks and violence. Secondly, information about political standpoints and reactions. Focus was introduced by providing the focus group first a text about the Irish terrorist movement IRA, also containing political information. Subjects were requested to compare the political situation in Northern-Ireland and Spain. Subjects in the non-focus (control) group received a neutral text about Spain without any specific reading instruction. Before and after reading the second ETA text the Cued Association task was presented.

The subjects read a list of 64 keywords twice and were instructed to write down next to each word the one, two or three words that came first to mind. The list contained 20 words that were important to the domain -based on a pilot study- but were not mentioned in the text, the other 44 keywords came from the text. The first answer to a cue (keyword) got a connection strength of 1, the second one mentioned 1/2, and the third 1/3. Answers that do not come from the list are ignored. We focused on the central parts of the representations, that is, the areas with the highest inner cohesion. The procedure to select these was based on (20) concepts with the highest in-degrees and out-degrees from and to other selected concepts respectively. To assess the amount of updating, that is the degree of change in the networks, we computed the similarity of the networks gathered before and after reading the second ETA text (see Ferstl & Kintsch, 1999). A high degree of similarity corresponds to a low degree of updating. In that case the network-before corresponds to a high degree with the network after reading the second ETA text. Also some textbase-oriented true/false recognition questions were presented at the end of the session.

Results

Because the size of the initial network structure of a reader -influenced by his/her background knowledge- could be a relevant variable to the extent of updating, next to the effect of focus on updating, or even interacting with it, we first determined the network size of each reader on the first cued association task. That is, the sum of the link strengths between concepts in the (first) matrix (see Ferstl & Kintsch, 1999). In other words, the more answers a reader produces consisting of keywords from the (first) list, the greater the network. Based on the mean score we distinguished subjects with an initial small and subjects with an initial large network size. A 2x2 analysis of variance (with focus and initial network size as factors) on the updating scores of the central parts of the networks only showed a significant interaction effect ($F(1, 37) = 4.02$, $p < .05$). For subjects with a small network there was less updating in the focus group compared to the non-focus group, while opposed to this, for subjects with a large network, focus showed a significant increase in updating compared to no focus. In a secondary analysis we examined what happened with the political and action concepts that were corrected in the text. We checked whether the embeddedness of political concepts (based on the in-degrees) in relation to the other, action concepts changed more in the focus group than in the non-focus group, since political concepts should be more in focus in the focus group. It appeared that for subjects with a small network the embeddedness of political concepts in other political concepts significantly decreased in the focus group compared to the non-focus group ($F(1, 17) = 7.09$, $p < .05$), while opposed to this, for subjects with a large network, focus leads to significantly less embedding of political concepts in action concepts compared to the non-focus group ($F(1, 17) = 5.27$, $p < .05$).There were no significant main or interaction effects of focus on the textbase-oriented questions.

Conclusions

Two conclusions can be drawn from these results. First, for subjects with small initial networks there is indeed some resistance to the updating of concepts relevant to the focus of reading. Subjects with initial small networks, reading with an unbiased reading goal, update to a higher degree and they embed the (correcting) political concepts to a greater extent. These results of a low degree of updating with a focus instruction correspond to what was mentioned in the introduction: Changes can be updated less easily with important information than with less important information, that is, information which is not in focus. Second, subjects with a initial large network, probably due to extensive background knowledge, update to a high degree when they read with focus. Furthermore, they show less embedding of political concepts in the action concepts. It seems like they construct two separate knowledge areas, one for political concepts and one for action concepts.

References

Chinn, C.A. & Brewer, W.F. (1993). The role of anomalous data in knowledge acquisition: A theoretical framework and implications for science instruction. Review of Educational Research, 63(1), 1-49.

Ferstl, E.C. & Kintsch, W. (1999). Learning from text: Structural knowledge assessment in the study of discourse comprehension. In H. van Oostendorp & S.R. Goldman (Eds.), The Construction of Mental Representations During Reading. Mahwah, NJ: Erlbaum.

Johnson, H.M. & Seifert, C.M. (1999). Modifying mental representations: Comprehending corrections. In H. van Oostendorp & S.R. Goldman (Eds.), The Construction of Mental Representations During Reading. Mahwah, NJ: Erlbaum.

van Oostendorp, H. (1996). Updating situation models derived from newspaper articles. Medien psychologie. Zeitschrift fur Individual- and Massenkommunikation, 8, 21-33.

van Oostendorp, H. & Bonebakker, C. (1999). Difficulties in updating mental representations during reading news reports. In H. van Oostendorp & S.R. Goldman (Eds.), The Construction of Mental Representations During Reading. Mahwah, NJ: Erlbaum.

Updating and the Reactivation of Situation-Model Information

R. A. Zwaan, C. J. Madden & R. A. Stanfield

Florida State University (U.S.A.)

zwaan@darwin.psy.fsu.edu

O'Brien et al., (1998) proposed that comprehenders do not update their information according to the here-and-now of the situation, as is claimed by theories of situation-model construction (e.g., Zwaan & Radvansky, 1998). That is, if later information qualifies earlier information in a text, this qualification does not prevent reactivation of the qualified (earlier) information. To test this hypothesis, O'Brien et al., used a clever technique. Their experimental stories came in three versions. In all versions, the beginning of the story described a personality or physical trait of a protagonist (e.g., being a vegetarian or being a junk-food lover). In the Consistent version, a later sentence described an action that was consistent with the trait. For example, a junk-food lover orders a cheeseburger with fries or an athletic young man rescues a little boy. In the inconsistent version, however, the action contradicts what is known about the protagonist, for example, a vegetarian orders a cheeseburger with fries or an octogenarian runs across the street and carries a boy to safety. Finally, and most importantly, in the Qualification condition, the initial description of the protagonist's trait is qualified in a later clause. Thus, the vegetarian is said to occasionally like to eat junk food and the octogenarian is said to be still fit.

O'Brien et al., rightly ascribe the prediction to the updating view that updated information should no longer be re-activated, as it is not part of the current situation. On the other hand, the resonance view proposed by O'Brien et al., predicts that incoming information sends activation to all of long-term memory, such that even information that was supposedly updated can be reactivated. Both views predict that the Inconsistent condition should yield longer reading times than the consistent condition, given that inconsistencies are difficult to integrate with the evolving situation model. However, only the resonance view predicts that the previous information will get reactivated and might thus cause some comprehension problems. However, given that there is not really a contradiction in the Qualification condition, the increase in reading times, relative to the Consistent condition, should be smaller than that produced by the Inconsistent condition. The updating view, on the other hand would claim that the qualified information replaces the previous information in the situation model, so that the Qualified condition should yield a pattern identical to that of the Consistent condition. In five experiments, O'Brien et al., find patterns that appear to be consistent with the resonance view and not with the updating view. In this presentation, we have three goals. We will first argue that some characteristics of O'Brien et al.'s materials do not allow for a proper test of the updating hypothesis. We then report two experiments that support the updating hypothesis, using the same paradigm as O'Brien et al. used, but with different materials. Finally, we will propose an account of our results that integrates the resonance mechanism with our model of situation-model construction (Zwaan & Radvansky, 1998). There are three characteristics of O'Brien et al.'s materials that are likely to have contributed to longer reading times in the Qualified than in the Consistent condition. Therefore, the O'Brien et al., findings cannot be used to properly evaluate the updating hypothesis. The first characteristic has to do with the question of what updating is. We argue that the manipulation used by O'Brien et al., does not involve exhaustive updating, such that the qualified information is still relevant to interpret the current situation. The second characteristic involves a potential confound in the materials which might have been partially responsible for the difference in reading times between the Qualified and Consistent conditions. The third characteristic also introduces a confound in that there was no equal billing for the original information and the qualification. The former was elaborated in much greater detail than the latter, which may have rendered the qualification less effective. In the presentation, we will discuss these characteristics of O'Brien et al's stimulus materials in detail.

We tested the updating and the resonance hypotheses using texts that met the following criteria: (1) the updating and updated information are mutually exclusive, i.e., if one holds, the other does not, (2) the updating information is not inconsistent with world knowledge, (3) the materials are free of a priming bias favoring one of the hypotheses, and (4) the updating information receives equal billing with the updated information. Consider the text in the Appendix. The first sentence states that Bobby's hammer is available for use in the Enablement condition, but not in the Disablement condition. Consequently, the sentence "The hammer was heavy for his young arm." creates a contradiction in the Disablement condition, but not in the Enablement condition. In the Re-enablement condition, the first sentence was identical to that in the Disablement condition, but it was followed by a re-enablement sentence, which, in this case states that Bobby found his hammer. Thus, if readers update their situation models, there should be no contradiction when they read the sentences about Bobby using the hammer.

We conducted two experiments using these materials. In Experiment 1a, the Disablement and Re-enablement texts were identical except for the fact that a re-enablement sentence was inserted in the Re-enablement condition right after the disablement sentence. In Experiment 1b, we tried to control for surface distance by inserting a filler sentence after the first sentence in the Enablement and Disablement conditions. Table 1 shows the results. As can be seen, they clearly support the here-and-now hypothesis. The fact that the instrument had been previously disabled did not elevate the reading times of the Re-enablement condition above those of the Enablement condition, whereas the Disablement sentence did yield reliably longer RTs than the other two conditions. We will provide an integrative discussion of these and O'Brien et al.'s results.

Table 1. Reading times (ms) for the critical sentences in Experiments 1a and 1b (standard deviations in parentheses)

	Enablement	Disablement	Re-enablement
Experiment 1A	2,074 (317)	2,281 (391)	2,096 (363)
Experiment 1B	2,236 (507)	2,413 (483)	2,228 (329)

Sample story: Title A BIRD HOUSE

Bobby took out a saw, but then remembered that he lost his hammer (Disable) / Bobby took out a hammer, but then remembered that he lost his saw (Enable) / After some searching, he found it in his father's toolshed (Re-enable) / After some searching, he found his father's tape measure (Filler, Exp.1b) / He also collected the lumber and paint he had bought (S1) / He had already selected an oak tree as the site for the birdhouse (S2) / He marked the boards and cut them out (S3) / Then, Bobby began pounding the boards together (Action) / The hammer was heavy for his young arm (Target sent).

The Effectiveness of Tutorial Dialog in an Automated Conversational Tutor

K. Link, V. Pomeroy, R. DiPaolo, S. Rajan, B. Klettke, L. Bautista, R. Kreuz, A. Graesser & The Tutoring Research Group

The University of Memphis (U.S.A.)

klink@memphis.edu

Many studies have demonstrated the effectiveness of tutoring in student learning (Cohen, Kulik, & Kulik, 1982). There are many possible reasons for this benefit, but one of them is undoubtedly the conversational interaction between the tutor and student (Graesser, Person, & Magliano, 1995). Tutorial dialog may allow for the specific remediation of knowledge deficits on the part of the student. In addition, this format engages the student because it is inherently interactive, as opposed to merely reading a text or listening to a lecture.

We have created an intelligent tutoring system, referred to as AutoTutor, that responds to student's natural language contributions in a manner that mimics the conversational strategies of a normal, unskilled human tutor (Graesser, Wiemer-Hastings, Wiemer-Hastings, Kreuz, & The Tutoring Research Group, 1999). AutoTutor's contributions are delivered via synthesized speech through an animated talking head with appropriate intonation and facial expressions (Person, Klettke, Link, Kreuz, & the TRG, 1999). The system is designed to help college students learn topics from a course in computer literacy.

Several modules in AutoTutor are used to understand the students' contributions. These include a word and punctuation segmenter, a part of speech tagger (Olde, Hoeffner, Chipman, Graesser, & the TRG, 1999), and a speech act classifier. In addition, latent semantic analysis (LSA; Landauer, Foltz, & Laham, 1998), a statistical technique used to represent world knowledge, is used by AutoTutor to determine the quality of the students' contributions (Graesser, Wiemer-Hastings, Wiemer-Hastings, Harter, Person, & the TRG, in press). Included in AutoTutor is a curriculum script that contains primarily concepts, facts, and question-answer pairs related to topics in computer literacy. The dialog move generator component of AutoTutor is composed of 15 fuzzy production rules that determine what, from the curriculum script, AutoTutor will say next. Depending on various aspects of the student's previous contribution and the preceding dialog in the tutoring session, AutoTutor will respond with one, or a combination of, the following: (1) pump, in which AutoTutor makes a request for the student to provide more information (e.g., "Anything else?"); (2) splice, in which AutoTutor inserts the correct answer to a question in response to a student's incorrect answer; (3) prompt, in which AutoTutor elicits a specific piece of information from the student (e.g., "ROM is a kind of _____?"); (4) hint, in which AutoTutor presents a fact or leading question, or reformulates the original question; (5) elaboration, in which AutoTutor contributes an important, but overlooked, piece of information; (6) feedback, in which AutoTutor provides an immediate evaluation of the student's last contribution that is either positive, negative, or neutral; or (7) summary, in which AutoTutor provides a brief synopsis of the answer or solution. For example, if the student's contributions have been minimal and if little of the topic currently under discussion has been covered, AutoTutor will offer a prompt.

The pedagogical effectiveness of AutoTutor was evaluated in an experiment by having 48 participants interact with the system. The participants were enrolled in an undergraduate computer literacy course and participated for course credit or for pay. The content of AutoTutor's curriculum script was developed from the textbook for this course. The topics on which participants were tutored in the experiment had been covered previously in their computer literacy course. Participants reread the chapter(s) on one topic and were tutored on a second topic. For a third

topic, participants reread the chapter(s) and were also tutored. Participants were then tested on these three topics: Computer hardware, Operating systems, and the Internet.

Two versions of the test were created, each of which consisted of three components: 18 shallow knowledge multiple choice questions, 12 deep reasoning multiple choice questions, and 72 recall questions presented in 18 Cloze passages. The shallow questions were selected from the test bank included with the textbook used in the computer literacy course. The deep questions were designed by the experimenters to get at causal chains (antecedents, consequences), goals, and purposes of procedures (e.g., "How can you best find information on the World Wide Web?"). Finally, the Cloze passages were created by removing key words and phrases from ideal answers to each question covered by AutoTutor in each of the three topics. Each component of the test assessed all three topics covered in the experiment.

On average, the participants interacted with AutoTutor for 63 min (SD = 17), and took an average of 144 turns (SD = 11). The mean performance on the test was 44% (SD = 13), and did not differ according to version (45% vs. 43%). A repeated measures analysis of variance showed that there were significant differences between the conditions ($F(2, 94) = 5.68$, $p < .01$). The text only condition resulted in the lowest performance on the test (M = 39%, SD = 15%). In contrast, the test scores on topics for which participants received only tutoring were higher (M = 47%, SD = 17%). The combined effect of tutoring and reading, however, was no better than the tutor only condition (M = 46%, SD = 16%). Paired comparisons showed that performance in the text only condition was significantly lower than the tutor only and the tutor and text conditions.

These results suggest that AutoTutor was an effective pedagogical partner. Specifically, the conditions that included AutoTutor resulted in 8% higher test scores than in the text only condition. Since the content in AutoTutor is functionally the same as the information in the textbook, we can conclude that the interactive nature of tutoring is responsible for this difference. These results are encouraging and suggest that intelligent tutoring systems require careful attention to conversational interaction. Other measures, such as participants, computer literacy course grades and self-assessments of computer literacy, will also be analyzed to evaluate the performance of AutoTutor more thoroughly.

References

Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. American Educational Research Journal, 19, 237-248.

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialog patterns in naturalistic one-on-one tutoring. Applied Cognitive Psychology, 9, 359-387.

Graesser, A. C., Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, D., Person, N., & The Tutoring Research Group. (in press). Using latent semantic analysis to evaluate the contributions of students in AutoTutor. Interactive Learning Environments.

Graesser, A. C., Wiemer-Hastings, K., Wiemer-Hastings, P., Kreuz, R., and The Tutoring Research Group. (1999). AutoTutor: A simulation of a human tutor. Journal of Cognitive Systems Research, 1, 35-51.

Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. Discourse Processes, 25, 259-284.

Olde, B. A., Hoeffner, J., Chipman, P., Graesser, A. C., & the Tutoring Research Group (1999). A connectionist model for part of speech tagging. In Proceedings of the 12th International Florida Artificial Intelligence Research Society Conference (pp. 172-176). Menlo Park, CA: AAAI.

Person, N. K., Klettke, B., Link, K., Kreuz, R. J., & the Tutoring Research Group (1999). The integration of affective responses into AutoTutor. In Proceedings of the International Workshop on Affect in Interactions(pp. 167-178). Siena, Italy: Springer.

The Dialog Advancer Network: A Mechanism for Improving AutoTutor's Conversational Skills

N. Person*, A. C. Graesser**, D. Harter** & The Tutoring Research Group**

*Rhodes College (U.S.A.)

**The University of Memphis (U.S.A.)

person@rhodes.edu

AutoTutor is an automated computer tutor that serves as a conversational partner with the student. AutoTutor is a working system that responds to Students' natural language contributions by simulating the dialog moves of normal (not expert) human tutors. Descriptions of how AutoTutor works have been discussed rather extensively in previous publications, and will therefore, not be provided in this brief proposal (see Graesser, Franklin, Wiemer-Hastings, & the TRG, 1998; Graesser, Wiemer-Hastings, Wiemer-Hastings, Harter, Person, & the TRG, in press; Hu, Graesser, and the TRG, 1998; Landauer & Dumais, 1997; McCauley, Gholson, Hu, Graesser, & the TRG, 1998; Olde, Hoeffner, Chipman, Graesser, & the TRG, 1999; Person, Graesser, Kreuz, Pomeroy, & the TRG, 1999; Person, Klettke, Link, Kreuz, & the TRG, 1999; Wiemer-Hastings, Graesser, Harter, & the TRG, 1998; Wiemer-Hastings, Wiemer-Hastings, & Graesser, 1999). The creation of AutoTutor was inspired by numerous studies that have systematically analyzed the collaborative discourse that occurs between human tutors and students (Fox, 1993; Graesser & Person, 1994; Graesser, Person, & Magliano, 1995; Hume, Michael, Rovick, & Evens, 1996; McArthur, Stasz, & Zmuidzinas, 1990; Merrill, Reiser, Ranney, & Trafton, 1992; Moore, 1995; Person & Graesser, 1999; Person, Graesser, Magliano, & Kreuz, 1994; Person, Kreuz, Zwaan, & Graesser, 1995; Putnam, 1987). One reoccurring finding in several of these studies is that human tutors rarely adhere to " ideal " tutoring models that are often integrated into intelligent tutoring systems. Instead, human tutors tend to rely on pedagogically effective strategies that are embedded within the conversational turns of the tutorial dialog.

AutoTutor's overall effectiveness as a tutor is contingent on him being an adequate conversational partner. Hence, many of our efforts have been directed toward implementing mechanisms that will enhance his conversational skills. One such mechanism is the Dialog Advancer Network (DAN). The DAN depicts AutoTutor's entire dialog move option space for any given student contribution (e.g., Assertion, Yes/No Question, Short Response). The DAN makes AutoTutor a more effective conversational partner in that it enables AutoTutor to: (1) adapt each dialog move to the previous turn of the student, and (2) indicate when the student has the floor to contribute. Both of these DAN functions are elaborated below.

Adapt each dialog move to the previous turn of the student. Coherence emerges in human conversations because participants generally try to adapt the content of their turns so that they are relevant (or linked) to the preceding conversational turn (Clark & Schaefer, 1987; Grice, 1975; Hobbs, 1979; McLaughlin, 1984; Nofsinger, 1991; Sacks, Schegloff, & Jefferson, 1978). This " turn-adaptation " process is somewhat problematic for AutoTutor because the content of his dialog moves is predetermined. That is, AutoTutor doesn't generate the content of his dialog moves on the fly but rather selects each dialog move from a scripted set of moves that is related to the tutoring topic being discussed. The DAN, however, enables AutoTutor to make quasi-adapted dialog moves that are relevant to the learner's preceding turn. As the DAN illustrates, AutoTutor assigns learner contributions to different speech act categories (e.g., Assertion, WH-Question, and Short Response) and then generates discourse markers, dialog moves, and answers to learner questions, accordingly. Thus, AutoTutor is able to sustain mixed-initiative dialog by acknowledging

student Questions (e.g., " What is RAM? ") and by responding to Short Responses (e.g., " I'm lost " or " I don't know ").

Indicate when the student has the floor to contribute Turn taking is an integral feature of the conversational process. To facilitate the turn-taking process in human-to-human conversations, speakers signal to listeners that they are relinquishing the floor (i.e., it is the Listener's turn to say something). However, human-to-computer conversations lack many of the subtle signals inherent to human conversations. When conversational agents like AutoTutor lack turn-taking signals, computer users (in our case, learners) often do not know when or if they are supposed to respond. For example, in previous versions of AutoTutor, learners were often confused after AutoTutor's Hint, Elaboration, and Prompt Response dialog moves. That is, the tutorial conversation would abruptly stop, and students would often ask lab assistants, " Am I supposed to say something now? " or " I'm not sure what to do ". This is illustrated in the following tutoring excerpt from a session about the CPU. AutoTutor says (Prompt dialog move): " See if you can get this, when the CPU executes programs, it uses RAM as . . . ? " Student types: memory. AutoTutor says (Positive feedback + Complete Prompt Response): " Good! As short-term memory storage. " Student says to lab assistant: " Am I supposed to type something in? "

In earlier versions of AutoTutor, the DAN was instrumental in helping us identify potential dead spots in the conversation (e.g., after Hints, Elaborations, Prompt Responses). We addressed this problem by adding more informative discourse markers and by adjusting the content of some of AutoTutor's dialog moves. As illustrated in the figure in the attachment, all pathways in AutoTutor's dialog move option space ultimately lead to the learner's next turn. In recent tutoring sessions with AutoTutor, learners appear to know when it is their turn to contribute. In the next version of AutoTutor, the DAN will include gestures and other paralinguistic signals (e.g., eye gaze) that facilitate the turn-taking process. In the presentation, we will provide additional excerpts from tutoring sessions with AutoTutor that illustrate how the DAN has enhanced AutoTutor's overall conversational capacities. We will also present data from a number of evaluative cycles in which AutoTutor's conversational skills were rated by knowledgeable judges. Some of the evaluative cycles occurred before the DAN was fully implemented.

References

Clark, H. H., & Schaefer, E. F. (1987). Collaborating on contributions to conversations. Language and Cognitive Processes, 2, 19-41.

Fox, B. (1993). The human tutorial dialog project. Hillsdale, NJ: Erlbaum.

Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. American Educational Research Journal, 31, 104-137.

Graesser, A. C., Person, N. K., & Magliano, J. P. (1995). Collaborative dialog patterns in naturalistic one-on-one tutoring. Applied Cognitive Psychology, 9, 359-387.

Graesser, A.C., Wiemer-Hastings, P., Wiemer-Hastings, K., Harter, D., Person, N., & the TRG (in press). Using latent semantic analysis to evaluate the contributions of students in AutoTutor. Interactive Learning Environments.

Hume, G. D., Michael, J. A., Rovick, A., & Evens, M. W. (1996). Hinting as a tactic in one-on-one tutoring. The Journal of the Learning Sciences, 5, 23-47.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychological Review.

McArthur, D., Stasz, C., & Zmuidzinas, M. (1990). Tutoring techniques in algebra. Cognition and Instruction, 7, 197-244.

McLaughlin, M. L. (1984). Conversation: How talk is organized. Beverly Hills, CA: Sage.

Merrill, D. C., Reiser, B. J., Ranney, M., & Trafton, J. G. (1992). Effective tutoring techniques: A comparison of human tutors and intelligent tutoring systems. The Journal of the Learning Sciences, 2, 277-305.

Moore, J.D. (1995). Participating in explanatory dialogues. Cambridge, MA: MIT Press.

Person, N. K., & Graesser, A. C. (1999). Evolution of discourse in cross-age tutoring. In A.M. O'Donnell and A. King (Eds.), Cognitive perspectives on peer learning (pp. 69-86). Mahwah, NJ: Erlbaum.

Person, N. K., Graesser, A. C., Magliano, J. P., & Kreuz, R. J. (1994). Inferring what the student knows in one-to-one tutoring: The role of student questions and answers. Learning and Individual Differences, 6, 205-29.

Person, N. K., Kreuz, R. J., Zwaan, R., & Graesser, A. C. (1995). Pragmatics and pedagogy: Conversational rules and politeness strategies may inhibit effective tutoring. Cognition and Instruction, 13, 161-188.

Putnam, R. T. (1987). Structuring and adjusting content for students: A study of live and simulated tutoring of addition. American Educational Research Journal, 24, 13-48.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1978). A simplest systematics for the organization of turn taking for conversation. In J Schenkein (Ed.), Studies in the organization of conversational interaction. New York: Academic Press.

Wiemer-Hastings, P., Wiemer-Hastings, K., & Graesser, A. C. (1999). Improving an intelligent tutor's comprehension of students with Latent Semantic Analysis. Artificial Intelligence in Education (pp. 535-542). Amsterdam: IOS Press.

Repetition in Discourse: A Linguistic Strategy Signifying Involvement in Children's Conversational Dialogues with a Literary Work

R. Horowitz & C. Cummings

The University of Texas -- San Antonio (U.S.A.)

rhorowitz@utsa.edu

Often repetition in oral and written discourse has been viewed as ineffective communication. Teachers admonish students for repetitive and incomplete utterances, which are also viewed as a mark of ineffective thinking. Tannen (1989) introduced a number of features characteristic of high involvement in dinner talk among adults. Repetition was one such characteristic. Weber (1995) hypothesized that repetition in child conversations might signal involvement with text. We will demonstrate, in the research to be presented, how repetition may be an important device in the creation of sustained dialogue and ideas and a sign of cognitive and affective involvement with literary text.

Rationale: There is little understanding about how involvement emerges or how it can be sustained during conversations about literary works. Julie of the Wolves (George, 1972), a winner of the prestigious Newberry Medal, was selected for this research because of its attention to culture and assimilation. This novel was chosen for its potential appeal to individuals in transition--including Mexican-American students who are exploring their cultural identities. It is a novel that immigrant children would find challenging because of the attention it gives to a child's cultural conflicts -- choosing between one's past traditions and the culture(s) of their new world. It addresses issues related to human survival in a markedly different world -- that of the animal kingdom.

Questions posed in the present research were: (1). What are the linguistic features that characterize high involvement in classroom conversations about text? (2). How does repetition function as a linguistic strategy of high involvement in classroom conversations? (3). What forms of repetition do students use? (4). How does the use of repetition evolve and change across four conversations?

Method

Subjects. Mexican-American, 5th graders, 11 years of age, in an urban classroom in South Texas participated in the research. Students were selected on the basis of their school schedule and willingness to participate in the project. They had a mean score of 6.8 ranging from 4.8 to 8.1 on the reading portion of the Metropolitan Achievement Test.

Procedures. The data was collected over a 2-week period in a departmentalized language arts block that met for 2 and 1/2 hours each day for 4 days. Pre-reading sources and activities were designed to build prior knowledge for the text. The classroom teacher videotaped the 4 conversations during the regular class time. The videotapes were transcribed by two teachers, a researcher and a community liaison, with each turn numbered in sequence. Two teachers scored the repetitions. There was ninety percent agreement in the categorizing of repetitions.

Results

Data Analysis. Repetitions used by students were categorized as: 1) self-repetitions of exact word or phrases that occurred within the same turn, 2) self-repetitions that were repeated from what students had read (not from what students had said), 3) self-repetitions of words, phrases or ideas that occurred across turns.

Findings. The children who participated in the study used repetition to serve a number of purposes as they explained and negotiated the meaning of, Julie of the Wolves. To summarize, the fifth graders used repetition to verify, for emphasis, and to persist, in their beliefs and opinions about the text. Functions of Repetition (examples provided in presentation):

1. Verification Repetition is a function of self-repetition. The speaker recalls or uses the text to verify something they have said to capture the listener's interest and show mastery of the novel.

   a.) Verification Text Language referred to the speaker's use of the author's language.

   b.) Verification Text Search was used more frequently by speakers and refers to a speaker's reference to a particular part of the text, where the speaker stated a page number, location of an idea.

2. Emphasis Repetition occurs when speakers repeat their own words, phrases, or sentences within a single turn for purpose of emphasizing an idea or event. There are two types of emphasis repetition:

   a.) Extension involves the adding of additional information-elaboration.

   b.) Dramatization refers to the speaker's use of language explicitly for the purpose of involving and drawing the audience in a dramatic way. Involvement is building and reflected by dramatization in the 3rd conversation.

3. Persistence Repetition occurs when a speaker repeats a word, phrase, statement or idea over a series of turns. There are two types of persistence repetition:

   a.) Turn Signal refers to a speaker's use of repetition to enter a conversation. It is a false start, which occurs during someone else's turn for the purpose of requesting the floor. This sets the audience up to expect something of importance to be said.

   b.) Stance occurs when a speaker has strong beliefs, takes a stand about those beliefs and as the conversation moves along, persists at returning to their position.

The turns in self-repetition increased significantly between conversations one and two ($p < .02$) and between conversations two and three ($p < .005$). Conversation four had a significantly higher percentage of repetitions than conversation one ($p < .005$). The turns in verification repetition increased between conversations two and three ($p < .02$) and decreased between conversations three and four ($p < .003$). The turns in emphasis repetition decreased significantly between conversations one and four ($p < .002$). The turns in persistence repetition increased significantly between conversations one and two ($p < .002$) and between conversations one and four ($p < .005$).

This study demonstrated that repetition is used by students in teacher-led conversations and functions as a linguistic feature marking and sustaining involvement. The students involved in the study used the functions of repetition to verify, emphasize, and persist while communicating their ideas to the other students. Engagement with the text increased across the four conversations. As

the conversations evolved, students had time to develop their beliefs about survival and acculturation. This was evidenced by a significant difference between the number of repetitions in conversations one and four. As the conversations continued, the students moved from a point where they were more text-focused (e.g., use of text verification was the highest in conversation three) to where they were engaged in emphasis and persistence of thinking about their own lives—conversations three and four. Conversation four shows a drop in text repetition, where students were able to step back from the text, and take an interpretive stance.

Conclusion

Most often discussions have been used to study levels of questioning and thinking. In the present research, discussion is viewed as the foundation for interpretation of a literary work. Discussion is a means for conceptualizing a text, engagement with a text, and for the creation of original ideas (Horowitz & Freeman, 1995; Horowitz, 1996, in press; Horowitz & Olson, in press). Our research suggests that extended conversations about text, over time, can help build the momentum that is needed for text involvement. Cultural and social considerations should also be a factor when looking at repetition in conversation. Repetition in conversation not only occurs in children of the United States but is also evident in other cultures of the world (Tannen 1989). The degree and type of repetition varies with the individual and culture of which one is a member. For example, African-American's use of repetition in call/response can be contrasted to the value put on silence by Athabaskan Indians (Tannen, 1989). In this study, the Mexican American student became emotionally attached to the cultural conflict of the main character, Julie, as she entered the world of the wolves. The study of repetition in small group conversations may explain the evolution of involvement and the stance speakers take when they become actively engaged with a text.

Accessability, Duration, and Modeling the Listener in Dialogue

E. G. Bard & M. P. Aylett

University of Edinburgh (Scotland)

ellen@ling.ed.ac.uk

Speakers are said to design their utterances to suit the needs of their listeners, insofar as those needs can be known (Ariel, 1990; Fowler & Housum, 1987; Gundel et al., 1993). Yet maintaining an incrementally updated model of what the listener needs to know is a considerable cognitive task. Tests of this claim do not often control the difference between listeners' and speakers' needs (see Keysar, 1997). Where such needs are known (Bard et al., 2000; Brown & Dell, 1987; Dell & Brown, 1991; Keysar, 1997), speakers' behavior seems to be more influenced by their own point of view than by the listener's. Dell and Brown (1987, 1991) propose that speech is initially generated on-line with little regard for the listener, unless the listener explicitly requests adjustments. A recent test of spontaneous speech failed to find even this accommodation. The length and intelligibility of words in referring expressions is sensitive to the status of the referent as New or Given (Fowler & Housum, 1987), with shorter, less clear tokens for the latter. Words from a large corpus of spontaneous dialogue (Anderson et al., 1991) proved to be shorter and less intelligible when referring to entities which were Given in so far as the speaker had heard them previously in reference to the same entity. Whether or not that speaker had produced that word token previously, or the listener had heard it, or either could see the object referred to (Bard et al., 2000) was unimportant. These results accord with Dell's suggestion that control of pronunciation is attributable to a kind of production priming which depends solely on the speaker's experience (Balota et al., 1989; Mitchell & Brown, 1988). The present paper asks whether either proposal holds for form of referring expression, where changes with repeated mention are usually abbreviations [the blacksmith's cottage….it] which can be assigned a place in a scale of referential accessability (Ariel, 1990; Gundel et al., 1993). By using all the references to landmarks in the original corpus, we can compare speakers' treatment of both pronunciation (via duration) and form of referring expression (via accessability) in situations where speaker and listener knowledge differ. In conjunction with current views of the language production process, Dell and Brown's proposals predict that referring expressions, whose structure is set at a planning stage before utterance begins, will tend to be egocentric in nature, like other early processes. Only if listeners provide negative feedback will a later mention be planned with the listener in mind. Bard et al., suggest, however, that is the duration of the process which permits computation of the listener's needs.

Method

We examined all expressions in the same route communication corpus which refer to landmarks, with the exception of those which were ambiguous as to accessability. In the HCRC Map Task Corpus (Anderson et al., 1991) instruction Giver's and Follower's maps matched only in alternate landmarks. Participants knew that their maps differed but not where or how. Each participant served as Instruction Giver for the same route to two different Followers and as Instruction Follower for two different routes. Digital stereo recordings were segmented at word boundaries, and coded for landmarks mentioned. Repeated references to landmarks using the same form of referring expression [the rift valley…the rift valley] or at least part of that expression [the valley] were compared for normalized duration (Bard et al., 2000). Those altering in form [the blacksmith's cottage….it] were classed by accessability on a scale running from 0 (cases with numeral or indefinite article + noun sequences) to 3 (pronoun, personal or relative).

Results

The experimental designs of Bard et al., (2000) were replicated for all suitable referring expressions in the corpus. In each case, changes in duration replicated: As long as the speaker had heard the words used with the same referent earlier, a second use was a shorter token. This finding held when the current listener had not heard the prior token ($F_2(1, 238) = 12.48$; $p < 0.0005$); whether the current speaker or the current listener had produced the earlier mention (mention: $F_2(1, 691) = 63.75$, $p < 0.0000$; mention x prior: n.s.) and whether or not the current speaker could see the object introduced by the current listener (mention: $F_2(1, 224) = 12.37$, $p < .0005$; mention x introducer: n.s.); and even when the listener had denied having the previously named object on his or her map (mention: $F_2(1, 193) = 9.45$, $p = .0024$; mention x denial n.s.). Changes in accessability differed. Accessability did sometimes take the listener into account: Accessability did not change where the listener had not heard the first mention (mention: $F_2(1, 115) < 1$). Yet accessability change was unaffected by intervening negative feedback (mention: $F_2(1, 135) = 23.33$, $p < 0.0001$; mention x denial: n.s.). However, accessability was also more sensitive to the speaker's own needs: Accessability increase across successive mentions was more limited when the current speaker lacked the object initially mentioned by the current listener (mention: $F_2(1, 138) = 24.67$, $p < 0.0001$; mention x sharing: $F_2(1, 138) = 6.48$, $p < .02$). Thus the results do not support Dell's prediction that earlier operations will be less sensitive to listener's needs, unless they follow pertinent feedback. Referring expression form was both more sensitive to listeners' needs and more sensitive to speakers'. This outcome is in line with the prediction that pronunciation and phrasal structure are composed by different mechanisms, with the latter available over longer-running processes than the former.

References

Anderson, A., Bader, M., Bard, E. G., Boyle, E., Doherty, G., et al., (1991). The H.C.R.C. Map Task Corpus. Language and Speech, 34, 351-366.

Ariel, M. (1990). Accessing noun-phrase antecedents. London : Routledge/Croom Helm.

Balota, D. A., Boland, J. E., & Shields, L. W. (1989). Priming in pronunciation: Beyond pattern-recognition and onset latency. Journal of Memory and Language, 28, 14-36.

Bard, E. G., Anderson, A., Sotillo, C., Aylett, M., Doherty-Sneddon, G., & Newlands, A. (2000). Controlling the intelligibility of referring expressions in dialogue. Journal of Memory and Language, 42, 1-22.

Brown, P., & Dell, G. (1987). Adapting production to comprehension -- the explicit mention of instruments. Cognitive Psychology, 19, 441-472.

Dell, G., & Brown, P. (1991). Mechanisms for listener-adaptation in language production: Limiting the role of the "model of the listener". In D. J. Napoli & J. A. Kegl (Eds.), Bridges between psychology and linguistics (pp. 111-222). Hillsdale: Erlbaum.

Fowler, C., & Housum, J. (1987). Talkers signalling of new and old words in speech and listeners perception and use of the distinction. Journal of Memory and Language, 26, 489-504.

Gundel, J.K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. Language, 69, 274-307.

Keysar, B. (1997). Unconfounding common ground. Discourse Processes, 24, 253-270.

McKoon, G., & Ratcliff, R. (1980). The comprehension processes and memory structures involved in anaphoric reference. Journal of Verbal Learning and Verbal Behavior, 19, 668-682.

Mitchell, D. B., and Brown, A. S. (1988). Persistent repetition priming in picture naming and its dissociation from recognition memory. Journal of Experimental Psychology: Learning Memory and Cognition, 14, 213-222

Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. Journal of Memory and Language, 37, 356-81.

Achieving Understanding in Multiparty Interactions

A. H. Anderson*, J. Mullin*, E. Katsavras*, R. McEwan*, E. Grattan* & P. Brundell**

*University of Glasgow (United Kingdom)

** University of Nottingham (United Kingdom)

anne@mcg.gla.ac.uk

Psycholinguistic studies of spoken discourse have been almost exclusively the study of two person dialogues. Research on computer-mediated interactions has more frequently examined the process of communication among small groups, although this has mainly been focused on turn taking behaviour. There is very little psycholinguistic research which compares the process and content of communication in dyads and small groups although the implications of the research of Herb Clark and his colleagues might suggest that achieving a shared understanding or " common ground " would be more difficult where more than two people are involved.

In a recent study (Anderson et al., 1999) we compared two and three party interactions supported by high quality videoconference links and compared these to comparable face-to-face interactions. We found that a similar level of task performance was achieved in the different types of interactions but that three party interactions, both face-to-face and mediated, required participants to expend more interactive effort, i.e. to use more talk to achieve their goal than in two party interactions. Here we explore what underlies these differences between 2 and 3 party interactions. In particular we explore to what extent the process of establishing common ground differs in multiparty interactions and the relative impact of increasing the size of the group from 2 to 3 with the introduction of multimedia communications technologies on these processes.

We used a collaborative problem solving task which we have used in previous research. The Map task has been found to be sensitive to the effects of communicative medium. Most importantly for the purposes of the present investigations it allows the researchers to control the knowledge available to each participant and thus to analyze the way in which participants resolve different kinds of knowledge mismatches and hence establish mutual knowledge during their interactions.

Method

One hundred and forty eight undergraduates were paid £5 for participating as members of a task pair or trio. Participants attempted two versions of the Map Task, either face-to-face or using videoconfernce links. They swapped between the instruction giver (IG) and the instruction follower (IF) role, on trials one and two. The maps which are schematic both/all show a start point and a number of named landmarks, some landmarks being common to all copies of the maps but others are shown on only 1 (or 2) of the maps. The participants' goal is for the instruction giver(s) to instruct the instruction follower to reproduce the route accurately. All are warned that some features of their maps differ.

In face-to-face conditions, paper maps were used and the three participants sat around a table with low screens preventing them from seeing one another's maps, although their faces were visible. In the VMC conditions, participants were first shown how the high quality videoconference links operated as they were all new to such technologies. In particular they were shown how images were captured and relayed, and how to use the mouse to draw on the onscreen maps. The

experimenters reminded participants that their collaborators were at a location three hundred miles away.

Results

Task outcome was evaluated by comparing the map drawn by the instruction follower with the original being described by the instruction giver(s). Performance did not differ significantly across the conditions. Maps were as accurately reproduced in two and three party VMC conditions and in three party face-to-face interactions.

The communication process was explored in analyses derived from transcriptions of the task interactions. Dialogues differed across conditions, $F_{(2, 54)}$, = 3.27, $p < 0.05$, with significantly longer interactions in both 3 party conditions. The mean numbers of words per conditions were: VMC 2 party 1199, VMC 3 party 1759, face-to-face 3 party 1649. A second length analysis was conducted using the total number of turns of speech. This showed a similar pattern of results. Two party VMC dialogues contained significantly fewer turns than 3 party VMC interactions, and difference between 2 party VMC and 3 party face-to-face conditions approached significance ($p = 0.08$). Three party VMC and face-to-face interactions again did not differ in length.

Analysis of Dialogue Content: Knowledge States

We wished to explore what caused the increased output in multiparty interactions. To explore these issues a new form of communication content coding was devised. This involved an exhaustive categorization of each turn of speech in terms of the speakers' and listeners' knowledge states about the topic (map landmark) being discussed.

Table 1. Knowledge State Categories

KS1:   Feature known to both/all

KS2:   Feature known only to IF

KS3:   Feature known only to one IG

KS4:   Feature known to IG1 & IF but not IG2

KS5:   Feature known to IG2 & IF but not IG1

KS6:   Feature known to IG1 & IG2 but not IF

Categories KS1, KS2, KS3 occur in both 2 and 3 party interactions while KS4, KS5 and KS6 only occur in three party interactions. To test the reliability of the coding scheme a Kappa coefficient was calculated on sample dialogues coded by both judges and showed good interjudge agreement, (kappa = 0.85). By comparing the lengths of interactions which concern directly comparable dialogue states i.e. KS1, 2 and 3 in the interactions we can assess whether there is a general increase in length in multiparty interactions. An ANOVA was conducted on these data which

showed that the length of these stretches did not differ according to the number of participants nor communication medium, (Fs < 1).

A separate analysis was carried out on the 3 party data, comparing the mean lengths of stretches of dialogue for all knowledge states. Knowledge state was a significant main effect, $F(5, 160) = 33.3, p < 0.001$). Post hoc tests showed that significantly lengthier discussions occur around knowledge states (KS4 & KS5) where there is a mismatch of knowledge between the two instructions givers. There was no difference between face-to-face and mediated interactions ($F < 1$).

Discussion

Anderson et al., (1999) showed that although two and three party interactions achieve equal levels of task success on a standard problem solving task, three party interactions both face-to-face and videomediated require more interactive work to achieve this outcome. The data clearly indicate however, that it is the complexities of multiparty interaction that seem to produce these lengthier interactions not the addition of technological support. In this paper we explored why three party interactions differed from those involving only two participants. The analyses of knowledge state clearly indicate that a key factor is the extra talk needed to align divergent knowledge states. Longer multiparty interactions arose from specific points in the task, notably from the difficulties in establishing common ground where knowledge is initially distributed among two of the three participants. Establishing common ground is significantly more difficult in multiparty interactions whether face-to-face or mediated by technology. Clark (1996) describes various possible representations for common ground as the basis for joint activities. Clark proposes that people operate in terms of a " principle of justification " that is they take a proposition to be common ground in a community only when they believe that they have a proper shared basis for the proposition in that community. Clark points out that the implications of this principle are that individuals expend considerable effort to find a shared basis for their common ground which in turns has impacts upon their language use. The data we have presented in this paper provide support for this position. The map task participants seem to be drawn by something like the principle of justification to extend this common ground throughout their temporary community. As this community enlarges from 2 to 3 participants this requires more collaborative effort. Thecommunication analyses suggest that the multimedia technology did not lead to any less of a sense of a language community. The challenge for future research is to understand more about the subtle characteristics of multiparty interaction, both face-to-face and mediated.

Psychology, Discourse and Ideology

T. A. van Dijk

University of Amsterdam (The Netherlands)

Universitat Pompeu Fabra, Barcelona (Spain)

teun@hum.uva.nl

This paper explores the role of ideology in discourse comprehension and production within the more general framework of a multidisciplinary approach to the study of the relations between discourse and ideology. The more specific aims of this paper are to examine the following aspects of this general theory: (a) The general nature of ideology within a more general theory of social cognition; (b) The structure of ideological representations in social memory; (c) The relations between ideologies and knowledge; (d) The ways ideologies influence mental models of events and context; (e) The strategic 'ideological' processing of discourse properties in comprehension and production.

A. Ideology and social cognition

A multidisciplinary theory of ideology assumes that ideologies have both social and cognitive properties. Whereas the social dimensions are defined in terms of groups, group relations, group interests and power, at the macrolevel, and in terms of many forms of ideological practices at the microlevel, the cognitive dimension of ideologies is defined in terms of representations in social memory and its role in the processing of discourse and other forms of social action and interaction. In other words, ideologies are forms of socially shared cognition among others, such as social knowledge, attitudes, norms and values. One of the theoretical problems it shares with these other forms of social cognition is what exactly it means that social representations are "shared" by the members of a group, and whether such sharing is more or less the same as the way sociocultural knowledge or the grammar of language are shared by social collectivities. Does this imply, for instance, that these representations need to be strictly the same, or only more or less comparable? Another theoretical problem is the internal architecture of the social mind, and how exactly ideologies are related to other socially shared representations.

B. The structure of ideologies

Any cognitive theory of ideology needs to spell out the structure of the mental representations that define an ideology. Traditionally, such an account would be given in terms of some kind of schema, that is, an ordered set of fixed categories used by social members (language users) to organize the ideological representations of the groups they are members of, or with which they identify themselves. Such a schema would explain the efficient acquisition, usage and change of ideologies in everyday social interaction (including discourse) within and between ideological groups. Comparisons with the possible structures of other forms of social representations, such as knowledge or attitudes, do not readily suggest a format for such ideological schemas. For instance, they are not like scripts, or like narrative and argumentative schemas. It is assumed that the categories of ideological schemas need to reflect the basic dimensions of social groups, such as (a)

membership criteria (who 'belongs' to Us, and who does not), (b) typical activities (c) goals (d) norms and values (e) social position (relations with other groups), and (f) basic resources.

C. Ideologies and knowledge

One of the classical epistemological problems in the (theory of the) social sciences has been the relationship between ideology and knowledge. Indeed, whereas knowledge is traditionally defined as 'true belief', ideological is usually negatively defined in terms of biased, partisan, distorted or otherwise misguided belief. Although this distinction may be further defined both in cognitive and social terms, our approach will not follow this evaluative approach. Indeed, ideologies may be very "true" and group knowledge may very well be "biased", depending on whose position or perspective is taken in their definition and evaluation. In our approach ideologies are not defined in terms of true or false beliefs, but in terms of their hierarchical relationship with respect to other forms of social cognition. Indeed, it is assumed that ideologies form the "axiomatic" basis of all forms of group cognition. This may mean that also the specific (and even specialized) knowledge of specific groups -- such as scholars, journalists, feminists, racists or ecologists -- that is, the shared beliefs held to be true within the group, may well be ideologically based. It will be argued however that this does not mean that *all* knowledge is ideologically based. Contrary to ideological group knowledge, for instance, we also distinguish socially shared (and generally presupposed) "common ground" knowledge, which is shared by a whole culture and hence does not differentiate between different ideological groups within the culture. The present paper will deal at some length with this problem of the cognitive nature of the difference between cultural knowledge and (ideologically based) group knowledge.

D. Ideologies and mental models

Ideologies influence mental models in various ways. First, by direct instantiation ("application") of general ideological propositions in the personal, individual evaluation of specific events. For instance, a general racist belief in the inferiority of Africans may be directly applied in all mental models featuring Africans as participants. This is not only true for mental models of events talked about, but also for mental models of communicative situations (so-called context models) in which members of a group may have ideologically based representations and evaluations of other participants. Secondly, ideologies may impinge on mental models indirectly, that is, through the instantiation of more specific forms of group knowledge and group attitudes, for instance in mental models of specific events related to immigration, abortion, nuclear energy or international policy.

In light of the theoretical framework sketched above, it will be assumed that producing or comprehending discourse in "ideological contexts" (that is, in social situations in which language users primarily participate *as* group members), may have itself a number of rather typical properties. Thus, it is often crucial that language users quickly and efficiently identify ideological group membership, typical activities, aims, norms, values, position and resources of current speakers or writers. That is, they must have strategies that allow them not only to understand a message (make a mental model for it), but also to understand its ideological "point" (make an adequate context model for it), namely as being conservative, liberal, feminist, racist, anti-racist, etc. This paper will examine some of these strategies of ideological discourse comprehension and production.

The Acquisition of Information Search Skills in 9 to 13 Year-Old Students

J-F. Rouet & C. Chollet

University of Poitiers and C.N.R.S. (France)

Jean-François.Rouet@mshs.univ-poitiers.fr

Studying often involves searching instructional documents for specific information. Locating information in documents requires the student to build appropriate search goals, to select relevant portions of the text, to evaluate content information, and to recycle the search process, updating the original goal structure if needed (Guthrie, 1988; Winne & Hadwin, 1998). It has been argued that information search is a specific cognitive skill distinct from reading comprehension, which may explain why it often proves difficult even for students with normal comprehension ability (Guthrie & Kirsch, 1987). Little is know about the development of search strategies in children. The ability to form goals, select, extract and evaluate information develops gradually between 4 and 10 years of age (Armbruster & Armstrong, 1993) and is related to learning and comprehension abilities (Raphael, 1984; Cataldo & Cornoldi, 1997). We investigated children's ability to search a junior encyclopedia in order to answer factual questions. We hypothesized that the number of search cycles and the amount of inferencing required to provide an answer may both affect children's performance quantitatively (e.g., search time) and qualitatively (e.g., search strategies).

Method

Participants. 36 children from grades 3 to 7 (9 to 13 year-old) participated as part of class activity.

Materials. A commercial paper encyclopedia on the topics of nature and wildlife, intended for a readership of 8 to 12 was selected from the existing educational literature. The book was organized into broad topics (e.g. deserts), which were subdivided in more specific topics, each presented on two side-by-side pages (e.g., the California Desert). Page design emphasized the integration of graphical and textual information, with many color drawings of landscapes, animals and plants. In addition the book included a table of contents and a subject index.

Tasks and procedure. Four search questions were designed. Two "simple" questions could be answered using a single passage in the book. Two "complex" questions required the comparison of two passages on different pages. In each category one question was textually explicit, i.e., the terms in the question could be matched directly with terms in the text. The other question was implicit, i.e., matching the question and the text required the child to generate a simple inference. For each question, relevant information could be found in the table of contents and in the index. The children participated individually in one half-hour session. The book was introduced to the child, and questions were asked one by one, both orally and on a separate piece of cardboard. The child was instructed to use the book so as to find an answer to the question as fast as they could. The child was allowed to re-read the question as many times as needed both before and while searching the text. No search method was explicitly suggested. Whenever the child remained inactive for a while, the experimented prompted comments or questions. In case of a difficulty the experimenter provided help according to a predefined sequence: (a) instruction to re-read the question, (b) explanation of keywords, (c) hints about alternate search procedures. The experimenter also provided feedback on the child's answers. In case of a wrong answer, the experimenter asked to the child to search further. The maximum amount of time devoted to one question was loosely set at 10 minutes. The experimenter transcribed the child's action (turning

pages, using table of contents or index, looking at headings, pictures, reading the question etc.) as well as the child's comments or questions.

Hypotheses. We expected that both complex and implicit questions would be harder to answer, especially for younger children. In addition, we expected that spontaneous strategies would evolve from simple page turning to top-down selection through the table of contents, index and page heading.

Results

Search time and correct answers. Search time was analyzed using a three-way ANOVA. Search time decreased with age ($F(2, 33) = 12.09$, $p < .01$), and increased for complex and implicit questions ($F(1, 33) = 11.83$, $p < .01$ and $F(1, 33) = 33.14$, $p < .01$, respectively). Furthermore, we observed an age by impliciteness interaction: Implicit questions took much more time than explicit questions for 9 years old and 11 years old, but not for 13 years old. Complex questions only took more time for 11 years old. The percentage of spontaneous correct answers (i.e., correct answers provided without any help from the experimenter) increased from 46% to 81% and 83% in 9, 11 and 13 years old, respectively. Implicit questions were the most difficult to solve for 9 year-old, whereas complex questions were the most difficult to solve for 11 years old.

Search strategies. Three types of strategies were observed. Flipping through pages, looking for a keyword or a picture was 9 years-old's most frequent behavior. Searching the table of contents was observed in 11 and 13 years old, while searching the index was the dominant strategy only in 13 years old.

Discussion

Searching an encyclopedia to answer factual questions was far from a trivial task for most of the participants. Many of them tended to forget the question, to lack an explicit strategy, or to ignore structuring devices such as the table of contents or the index. The ability to search more quickly and efficiently seems to develop gradually between the ages of 9 and 13. However, some 13 years old students still rely on rudimentary strategies such as flipping through pages and reading everything from top to bottom. Our data suggest that 9 years old students are mostly concerned with reading words, accessing word meanings, and confronting words in the question with words in the text. Hence their frequent failure when the text and the question do not exactly match. Thus, search efficiency in 9 years old seems strongly constrained by vocabulary and reading fluency. Remembering the question while searching is another problem, especially in 9 and 11 years old, which suggests that the development of working memory is another prominent cognitive factor of search efficiency in children (and possibly later on). Finally, students' knowledge of structuring devices (headings, page layout, standard superstructures, table of contents…) may also strongly influence their ability to achieve efficient search strategies. Our current research efforts aim at assessing the respective importance of working memory, reading fluency and comprehension ability on various measures of search efficiency in 9 to 13 years old students. Also of interest is the issue of whether and how children in the elementary grades can be taught more efficient search strategies.

References

Cataldo, M.G. & Cornoldi, C. (1998). Self-monitoring in poor and good reading comprehenders and their use of strategy. British Journal of Developmental Psychology, 16, 155-165.

Guthrie, J.T. (1988). Locating information in documents: examination of a cognitive model. Reading Research Quarterly, 23, 178-199.

Guthrie, J.T. & Kirsch, I. (1987). Distinctions between reading comprehension and locating information in text. Journal of Educational Psychology, 79, 210-228.

Raphael, T.E. (1984). Teaching learners about sources of information for answering comprehension questions. Journal of Reading, January, 303-311.

Winne, P.H. & Hadwin, A.F. (1998). Studying as self-regulated learning. In D. Hacker, J. Dunlowsky & A. Graesser (Eds.) Metacognition in educational theory and practice (pp. 277-304). Mahwah, NJ: Erlbaum.

LSA in the Classroom: Automatic Feedback for Learning Summarizing Skills

E. Kintsch & D. Steinhart

University of Colorado at Boulder (U.S.A.)

ekintsch@psych.colorado.edu

LSA is a theory of meaning and knowledge representation developed by Landauer and his colleagues (see, for example, Landauer & Dumais, Psychological Review, 1997; Landauer, Foltz, & Laham, [Special Issue], Discourse Processes, 1998) that has rich potential for educational as well as other applications. Our presentation will focus on one of these applications, known as *Summary Street*, which is intended to help students learn how to summarize instructional materials. We will briefly discuss the educational rationale for this tool and describe the kind of feedback it provides. However, the body of the talk will concern the results of recent classroom trials with the tool.

Our goal in designing a computer tool to support the writing of summaries stems from the curricular needs expressed by two 6th-grade teachers with whom we have been collaborating. Their instruction emphasizes summarization as a means for students to deepen their understanding of complex new material, as well as to improve their writing and revising skills. The problem addressed by *Summary Street* is the need to provide students with enough practice in extended writing to accomplish these goals without producing an overwhelming amount of work for teachers in evaluating successive drafts of students' writing. Using LSA's knowledge analysis capability, feedback on the content of students' writing can be automatically delivered, allowing students to do much more revising without human guidance than has previously been possible. *Summary Street*, however, is not intended to replace the advice of human teachers, but rather to enhance their effectiveness. The feedback delivered by *Summary Street* is mainly in the form of a graphic display indicating to what extent each subtopic section of the source text has been covered by a student's summary. Missing topics are signaled with a link to the appropriate information in the text. The system also indicates whether the summary meets the teacher's prescribed length requirement, and it marks misspellings and flags sentences that may contain irrelevant or redundant content.

Method

We used *Summary Street* in a series of classroom trials that formed part of the regular instructional curriculum during the spring and fall 1999 academic semesters. The three curricular units that involved summarization activities were about ancient Meso-American cultures, the human circulatory system, and renewable and non-renewable sources of energy. Two classes of 6th-grade students (52 total) participated in these trials during which they wrote and revised their summaries either using the summarization tool or traditional means (pen and paper or word processor). The source texts varied from about two to four pages in length, and summaries were supposed to be about a quarter or less of the original length (c. 150-300 words maximum). The purpose of the two trials that took place in spring 1999 (Meso-American and Heart-Lungs texts) was mainly to try out the system rather than to conduct a formal empirical test. Nevertheless, one intriguing result was consistently obtained in these trials: Namely, even though summary grades did not differ across conditions in these two trials, *Summary Street* summaries of certain texts received significantly higher scores than traditionally written ones. In both trials these were the texts that both teachers and students rated as more difficult, which was also reflected in the students' grades. The trial in fall 1999 on energy sources took place under more controlled circumstances, in order to formally evaluate the system. During this trial students revised two

summaries either guided by the full content feedback from *Summary Street* (Feedback condition) or using a similar interface that only provided feedback on length and spelling (No Feedback condition). Conditions were switched for the second summary written a week later. This procedure allowed us to track the time students spent working on their summaries, as well as successive drafts of students' work, and to look for possible transfer effects.

Results

The two classroom teachers, who were blind to students' identity and condition, scored the summaries (a) for adequacy of content coverage of each subtopic section; and (b) for its overall, holistic quality. Our main results are as follows:

Time on task: Students who received feedback on content coverage from *Summary Street* spent more than twice as long working on their summaries than those did who received no feedback, 72 min. vs. 33 min., respectively.

Content scores: Averaged content scores (based on a scale from 0-2-points) for *Summary Street* summaries were significantly higher than those in the No Feedback condition: 1.29 points on average vs. 1.01 points, respectively.

Overall Quality: The teachers used a 5-point scale to rate the summaries' holistic quality, taking into account factors such as coherence, style, mechanics in addition to content. These scores were also significantly higher in the case of *Summary Street* summaries (3.19 points on average) than for summaries written without content feedback (2.87 points).

Text Difficulty: As in the previous trials, the benefits of *Summary Street* feedback were greatest for the most difficult texts and low to non-existent for the easier ones. This result suggests that the support offered by the tool is most effective when students are struggling with difficult text and task requirements (e.g., tight length constraints).

Individual Differences: We also found that the system was more effective for some students than others. Namely, moderate achievers seemed to benefit most from the feedback they received, whereas students whom the teachers considered high achievers apparently did not need the supportive feedback to write good summaries. Low achieving students, in contrast, could not profit as much from the feedback with texts that may have been too difficult for them. Nevertheless, these students were observed to work much longer and with greater interest using *Summary Street* than with most writing tasks.

Discussion

Overall, these results offer strong support for the assumptions underlying the design of *Summary Street*: By automatically supplying students with feedback on the content adequacy of their writing, they can receive much more writing practice than is normally the case. This is because students can themselves assess this basic component of writing quality, rather than having to rely on a teacher's availability. Educators often complain that students rarely revise their writing, probably because throughout their academic training they have few opportunities to do so. However, we find that the feedback itself, together with the simple, but engaging environment in which it is delivered, is remarkably successful at keeping students fully engaged through repeated cycles of revising over long periods of time.

The Use of Narrative in Argumentation

J. F. Voss & J. van Dyke

University of Pittsburgh (U.S.A.)

Voss@vms.cis.pitt.edu

In recent years the study of argumentation text and of narrative have been of particular interest, both within psychology and in other disciplines. A distinction made about human thinking related to these topics is that one type of thought is of a more "logical" nature and includes the mental activity involved in mathematics and the various forms of logic tasks, and the other may be regarded as narrative, a sequential story telling type of endeavor (cf. Bruner, 1986). While this distinction is reasonable, and may be supported by lines of evidence, the idea that these two realms of study are in fact completely separated, seems to push the issue too far. Indeed, it was Aristotle who pointed out that the narrative could be used in argumentation, especially as it relates to courtroom situations. We have been investigating the issue of how narrative may be used in relation to argument, and what effects using narrative may have upon judgment. In a previously published initial study, participants were presented with an account of a hypothetical murder case, as presented to them by a prosecuting attorney. The attorney's account was a narrative of the committing of the crime, as the attorney interpreted it. We manipulated four narrative conditions, one being a baseline standard narrative, and other three being degraded in relation to the narrative's coherence and chronology (which cannot readily be separated), the narrative's causality, and the narrative's completeness (The completeness condition had a problem and is not discussed further). Narrative condition and four different texts were counterbalanced so that each participant served once in each condition and one time with each text. The basic idea of the study was that when the prosecuting attorney provides the account of the crime, the narrative may be divided into two components, the basic evidence, that is, the fingerprints found on the baseball bat, who was seen with the bat, etc. and the parts of the narrative that are not the hard evidence but consist of other parts of the narrative. What we then did is hold the basic evidence constant, but varied the narrative conditions. Participants then made two types of judgments after hearing and reading the respective narrative. They first judged the guilt of the accused on a 1-10 scale, and then they made judgments about the quality and persuasiveness of the narrative.

The hypothesis was that in those narratives that were degraded, the guilt judgments would be lower than in the baseline narrative. Moreover, the ratings of the quality of those narratives also should be lower than for the baseline narrative. These results were obtained. The primary significance of the results is that the quality of the narrative was shown to produce lower guilt ratings, even though the basic evidence was constant for all narratives. Looking at these results from the perspective of argumentation, the following analysis may be shown. The prosecuting attorney makes the claim that X, the accused, is guilty. To support this claim, the prosecutor provides a supporting narrative. But when the narrative quality is degraded, the attorney's argument is not as well as accepted as when it is not degraded. (Although not mentioned previously, the characteristics of what constitutes a good narrative were obtained from two studies, one in which historians provided attributes of a good narrative and one in which the attributes were obtained from the study of jury decision making.)

The two studies to be reported constitute a continuation of the above research. In one, the coherence/chronolgy condition is investigated in greater detail and in the second the causality degradation is studied. In addition, the possible effect of emotional textual statements is studied, the emotion pertaining to the crime or to the victim. The evidence presented in the first study was strong enough to produce guilt ratings of about 7 on the guilty, not guilty 10-point scale, but is would not be regarded as quite trong. A question then posed by the results of the first study is

whether if you had good and poor evidence conditions, a good narrative could compensate for poor guilt ratings. A 2x2 study of quality of evidence and coherence quality was conducted. Coherence was manipulated by having a baseline text, with either good of poor evidence, and a text in which the baseline sentence order was scrambled, the manipulation used in the first study.

The results indicated that good evidence yielded significantly higher guilt ratings than poor evidence, with good evidence also yielding higher judgments of text quality and convincingness. In addition, males made more extreme guilt ratings than females. However, while the narrative manipulation influenced judgments of text quality and convincingness, guilt judgments were not significantly different. Clearly, a good narrative did not compensate for poor evidence.

In the other study causality was manipulated by including statements that made the evidence more probabilistic and less certain as "The fingerprints of the bat were Norman's" versus "The fingerprints on the bat were probably Norman's" In a second condition, the same type of manipulation took place, but if involved subject matter irrelevant to the crime. This was done to determine whether making irrelevant information more probabilistic could have an effect on guilt, a type of general effect. In addition, there was a baseline condition and an emotion condition in which the emotion pertained to the crime and another condition in which it pertained to the victim. Basically, the results indicated that the crime-related causal degrading influenced the guilt judgment and the textual judgments. The irrelevant causal condition has a weak effect. Emotion related to the crime did not influence guilt judgments, but it did influence the text judgments, for women only. The emotion related to the victim influenced both the guilt judgments and text judgments, but only for women.

The results suggest that evidence is critical for both guilt judgments and the judgments of narrative quality, that evidence overrides coherence as long as there is an intelligible level of coherence and the evidence is strong, that emotional contests influence judgments of text quality for women, and for women guilt judgments were influenced by victim-based emotion statements.

The Understanding Of Nominal Metaphor

C. Tijus, B. Pudelko, E. Hamilton, & D. Legros

University of Paris VIII (France)

tijus@univ-paris8.fr

The objective of this research is to explain how the choice between different figurative meanings is carried out. To this end, we place the accent on the role of context in everyday situations of metaphor comprehension. We consider that in ecological situations of communication, the context comprises not only knowledge about the world but also new information provided by the situation in which the topic of the metaphor appears. Metaphor disambiguation is based on information furnished by the context which constrains the choice of meanings for the metaphor, and this, even in the case of conventional metaphors. Most of the cognitive models of metaphor understanding have adopted the approach according to which metaphor is an implicit comparison (Gentner & Wolff, 1997 ; Ortony, 1979 ; Tourangeau & Rips, 1991). From this point of view, understanding a metaphor "X (topic) is Y (vehicle)" consists in converting it into a simile "X (the topic) is like Y (the vehicle)" The mechanism posited by the comparison models is a mechanism of property matching. This is why these models are confronted with the problem of measuring the similarity of properties as well as with the problem of calculating the distance between properties which makes a simile literal or metaphoric. More recently, an alternative categorization approach has been proposed by Glucksberg et al., (1997) and Way (1991). According to Glucksberg's class inclusion model, a metaphoric statement of the type "X is Y" is solved by looking for the category, represented by the term Y, which furnishes properties that are potentially relevant for the topic X. The disadvantage of this model is that there is no mechanism for explaining how the relevant properties are selected. From our point of view, this aspect must be adressed if we are to explain how metaphors are disambiguated, i.e. to explain why one property is chosen and not others.

Our general hypothesis is that metaphor understanding consists of including the topic in the category of the vehicle and attributing to it the properties of that category that are compatible with what we already know about the topic. Legros, Tijus et Pudelko (1998) have shown that when the context provides specific information about the topic, the selection of vehicle features is compatible with this information. We are assuming that interpretation is constructed on-line and that knowledge about the topic intervenes at an early stage in processing by constraining the selection of features.

Method

The experiment we report here aims at studying the way in which different contexts providing information about the topic activate features of the vehicle. We opted for presenting 14 metaphors in mini-scenarios in order to make the laboratory situation more like usual conversational situations in which animal metaphors are used. The scenarios were of three types: designed (i) to be compatible with conventional interpretations, (ii) to be compatible with one of the actual features of the animal or (iii) to be neutral, i.e. not particularly compatible with conventional or with novel interpretations (Table 1).

26 participants were first presented with the context sentence, then with the metaphor, and finally with an interpretation sentence. We tested property transfer by alternately presenting three different interpretation sentences of a given metaphor ("Does this mean that Guy is… funny? agile? likes music?"). Participants had to say whether or not they agreed. We predicted that (i)

when the interpretation sentence is not compatible with the context, the agreement score would be low; (ii) when the interpretation sentence is both compatible with the context and with a real or conventional property of the animal in question, the agreement score would be high; and, (iii) when the context is neutral, the rates of agreement would be higher in this case for the conventional interpretation which predominates.

Results

Table 1. Percentages of "Yes" responses and Reaction Times of "Yes" responses. The 3x3 experimental conditions where, for each metaphor 3 context phrases describing a person in terms of (i) an actual animal trait (e.g. "likes to climb"), (ii) a conventional animal trait ("likes to joke") or (iii) an unassociated trait (e.g. "often goes to concerts") were crossed with 3 target phrases in which a trait is attributed to a person (i) an actual animal trait (e.g. "agile"), (ii) a conventional animal trait (e.g. "funny"), or (iii) a trait not associated with an animal (e.g. "likes music"). Each cell represents 364 data; 26 participants responded to the 9 conditions for each metaphor.

Context sentences	Target-sentences		
	actual animal Guy is agile	unassociated Guy likes music	conventional Guy is funny
actual animal "Guy likes to climb"	86% (2.18 s)	5% (3.21 s)	41% (2.65 s)
unassociated "Guy goes to concerts"	28% (4.28 s)	35% (3.53 s)	41% (3.36 s)
conventional "Guy likes to joke"	36% (3.04 s)	6% (2.46 s)	91% (2.10 s)

Discussion and conclusion

Our results show that the context in which the topic appears constrains the selection of the relevant trait to be transferred from the vehicle. Indeed, context can determine the selection of either conventional or novel traits. No effect was observed for kind of trait as long as both context sentences and interpretive sentences were compatible, in either case the metaphor was disambiguated quickly. Neither was kind of trait effect observed when the context sentence was incompatible with the target sentence.

How can the effect of initial knowledge about the topic in the construction of meaning for metaphor be explained, and in particular, how are elements of this knowledge selected? The CADS-T model proposes that the mental representation of situations described using language is composed of category networks which are formed by factorising the properties activated by the context as well as through a property adjustment process. CADS-T enables attributing a meaning to the topic through selecting a vehicle property as a function of the constraints provided by the topic's context. As we see it, the comparison and the categorization approaches are not antagonistic but

complementary: understanding a metaphor consists both in (i) including the topic in the vehicle category and (ii) in selecting the properties that are compatible with the context.

References

Gentner, D., & Wolff, P. (1997). Alignement in the processing of metaphor. Journal of Memory and Language. 37, 331-355.

Glucksberg, S., McGlone M.S., & Manfredi, D. (1997). Property attribution in metaphor comprehension. Journal of Memory and Language, 36, 50-67.

Legros, D., Tijus, C., & Pudelko, B. (1998). The effect of knowledge about topic on metaphorical meaning. In Proceedings of the 20th Annual Conference of the Cognitive Science Society. p.1238. Mahwah, NJ: Lawrence Erlbaum Associates.

Ortony, A. (1979). Beyond literal similarity. Psychological Rewiev, 86 (3), 161-179.

Searle, J. (1970). Expression and Meaning. London, Cambridge University Press.

Tijus, C., & Moulin, F. (1997). L'assignation de signification étudiée à partir de textes d'histoires drôles. L'Année Psychologique, 97, 33-75.

Tourangeau, R., & Rips, L., (1991). Interpreting and evaluating metaphors. Journal of Memory and Language , 30, 452-472.

Way, E.C. (1991). Knowledge representation and metaphor. Dordrecht: Kluver Academic Publishers.

The Costs and Benefits of Metaphor

I. A. Noveck* & M. Bianco**

*Institute of Cognitive Sciences, Bron (France)

**University of Grenoble (France)

noveck@poly.polytechnique.fr

In this paper, we briefly describe Relevance Theory (Sperber & Wilson, 1995) and a cost/benefit explanation of metaphors that the theory inspires. We argue that apt, unanticipated metaphors require extra cognitive effort to process. We present two developmental studies that buttress our claims based on a paradigm from Gibbs (1990). Our ultimate aim is to show how Relevance Theory can do a good job of accounting for psycholinguistic findings on metaphor.

Relevance Theory views inference-making as a constant feature of communication geared towards gathering in (and sharing) one's intentions. Essential to Relevance Theory is the claim that, in processing any utterance, people endeavor to draw out as many cognitive effects (i.e. benefits) as possible for the least effort (i.e. cost). Two features of Relevance Theory are crucial when describing metaphors (see Sperber & Wilson, 1986): A) Utterances need not be literally true in order for a listener to draw implications effectively and; B) A metaphoric utterance is likely to carry more information than a more-literal equivalent. To illustrate both features, consider a scenario in which a swimming instructor says "you are a tadpole" to a 5-year-old. The utterance a) is not literally true while effectively conveying information from teacher to student and it; b) goes further than a literal equivalent (something like "you are a young child doing a frog kick"); at least, the instructor's metaphor is arguably endearing whereas the literal equivalent is not.

The second feature is of particular interest. Relevance Theory essentially argues that the metaphor prompts multi-tasking. In the tadpole example, the metaphor is making reference to the swimming student plus it is describing something about him and transmitting affection; a neutral expression like "you are a child" in the same context would be doing the reference portion only (and would not seem terribly informative). This analysis shows that there are (at minimum) two components to a full appreciation of metaphor: Understanding what the metaphor is referring to and understanding the interlocutor's intention in using it. Being that it is difficult to imagine that multiple cognitive effects associated with metaphor come cost-free, we make two claims. Our primary claim is that it should not be surprising to find that -- all other things being equal -- metaphors are costly to process when compared to literal controls. Our auxiliary claim is that the extra costs that come with an apt metaphor ought to be commensurate with extra benefits. That is, one should be able to detect that apt metaphors have benefits when compared to less-costly, non-figurative equivalents.

To address these issues more directly, we investigate the development of metaphor comprehension because differences in effort ought to be even clearer among less proficient readers. We adopted a paradigm from Gibbs (1990), in which he presented participants with seven lines of a story before presenting one of three different concluding sentences that varied their form of reference (metaphoric, synonymous, and new). We prepared stories in French having metaphoric or synonymous references in the penultimate line and a follow-up question. For example, one story told of second-graders who went off to a swimming class. At one point during the class the teacher was interrupted by a phone call and upon returning he ordered all the children to the side of the pool, at which time they were sent to their lockers. The second to last line of this story was expressed either as:

1a.   All the toads to the side of the pool (metaphoric reference).

1b.   All the students to the side of the pool (synonymous reference).

In two experiments, stories were distributed so that each participant would see half the stories with a metaphoric reference and half with a synonymous reference.

Experiment 1

Two hundred and thirty children between the ages of 8 and 12 were presented 16 stories on paper. All the stories provided questions that directly asked about the referent and all the questions required a "yes" or "no" response. For example, with respect to the story above the question (presented while the text was available) was "Were the pupils the ones who went to the side of the pool?" Regardless of formulation (metaphoric or synonymous), the correct response was "yes." We found two revealing effects. First, rates of correct responses show that referential ability in general improves with age. Second, formulations with synonymous references are linked to rates of correct responses that are consistently about 7% higher than those with metaphoric ones until around 12 years of age. Among 12 years old, one sees the gap close (to 2.6%). This indicates that metaphors do come with a small risk of leading young readers astray but the risk appears to diminish with age.

Experiment 2

Fifty nine-years-old, 48 eleven-years-old, and 51 fourteen-years-old were presented 12 of the sixteen stories from Experiment 1. Forty adults were presented all sixteen stories from Experiment 1 plus filler items. In this experiment, each story was accompanied by one of three kinds of follow-up questions that justifies either a "yes" or "no" response: a) a question about a detail of the story, b) a general comprehension question, or c) a question like the one in the first experiment concerning the referent. Only one kind of question was presented after each story. The dependent measures were the reading times of the penultimate (metaphoric or synonymous) line and responses to the follow-up questions. Analyses concerned only those stories that were seen by both the children and adults. The reading times revealed that, at each age, sentences containing the metaphoric reference are read consistently more slowly than those containing the synonymous reference and that the gap closes with age, though never completely. The adult data confirm that a metaphoric reference in Gibbs's paradigm prompts a significant slow-down when compared to a synonymous control (a result that Gibbs, 1990, tended to disregard). The rates of correct responses confirm that the youngest children pay a small price in comprehension when they encounter a metaphoric reference. Adults' responses reveal that the metaphoric reference actually aids comprehension slightly. We remain cautious about our second finding because only the adults saw filler stories. Nevertheless, the pattern of results is consistent with the cost/benefit analysis that led to our initial predictions.

In conclusion, we argue that metaphors can be analyzed in terms of costs and benefits. Our findings show that a) compared to controls, metaphoric references consistently prompt longer reading times and b) in terms of comprehension, metaphoric references are sources of difficulty for younger children and sources of potential benefit for adults. We thus hope to show that a metaphoric reference is an imposition on a reader but that its potential for impact is linked with an ability to appreciate its intended meaning.

References

Gibbs, R. (1990). Comprehending figurative referential descriptions. Journal of Experimental Psychology: Learning, Memory and Cognition, 16(1), 56-66.

Gibbs, R. (1994). The Poetics of Mind: Figurative Thought, Language, and Understanding. New York: Cambridge University Press.

Sperber, D. & Wilson, D. (1986). Loose Talk. Proceedings of the Aristotlean Society, 86 (1985-1986), 153-171.

Sperber, D. and Wilson, D. (1995). Relevance: Communication and Cognition. Blackwell, Oxford.

Prosodic Correlates of Text Structure

H. J. N. den Ouden*, L. G. M. Noordman* & J. M. B. Terken**

*Tilburg University (The Netherlands)

**University of Technology, Eindhoven (The Netherlands)

J.N.denOuden@kub.nl

Authors use typographical means such as sentence and paragraph boundaries in order to structure their message. Our research deals with spoken texts. The main question in this research is whether prosody has a comparable function in conveying information about text structure in spoken texts as typography has in written texts. More concretely, we address the question whether and how prosodic features are correlated with text structural features. Ultimately we search for a model for mapping the text structure onto prosodic structure. The research presented in the current paper consists of two parts. The first part deals with the question whether a theory designed for the linguistic analysis of texts can be applied reliably for annotating text structure and the second part investigates the relation between the linguistic analyses of texts and their prosodic features.

In order to find prosodic characteristics of discourse structures, we need a description of the structures of texts. To avoid circularity, the identification of discourse units and the relations between them should be independently motivated and not depend on prosodic considerations. There are several ways to obtain discourse structures. One possibility is an empirical approach by asking a number of judges to determine paragraph boundaries in the text. The result of this approach is a text segmented into several units, between which the cohesion differs depending on the number of judges who determined a break as a boundary. Although this is an objective and scientifically sound method, it leaves people's intuitions about the grounds for segmentation implicit.

An alternative method is to employ linguistic theories that describe the structure of texts. The advantages of such a theoretical approach are twofold. First, it does not require a large number of judges. Second, the resulting text structures are not based on implicit intuitions, but are theoretically motivated. However, before such theories can be applied for our purpose, we need to assess whether they are reliable.

The theory we used in our research is Rhetorical Structure Theory (Mann & Thompson, 1988, henceforth RST). RST analyzes a text into a hierarchical structure with labeled relations between discourse segments and clusters of segments by defining explicitly the conditions to be fulfilled for each relation. To find out whether RST can be applied in a reliable way, we conducted a study to test the inter-judge reliability. We used two news reports, representing informative texts, and two commentaries, representing argumentative texts, all of which had been broadcast on Dutch radio. The texts were segmented into elementary units according to RST's definition of a clause. All texts had a length of about 30 segments.

Six expert RST analysts were asked to analyze the texts segmented into elementary units in terms of RST as reported in Mann and Thompson (1988). Although the texts had been spoken originally, the analysts had access to the written transcripts only. The output consisted of one hierarchical text structure per text per analyst. In order to compute the inter-judge reliability, quantitative representations of the hierarchical text structures were obtained in the following way. Each boundary in the hierarchical text structure of each analyst obtained a level number, starting at 1 from the topmost branching node downwards to the lowest branching nodes. The reliability was determined by computing Kendall's measure of concordance W between the scores of the individual

analysts for each text. The agreement between the analysts was .72, .73, .51 and .79 for the texts respectively. For all texts these reliability results were significant.

In addition, we compared the reliability for RST with that for the theory based on Grosz and Sidner (1986, henceforth G&S). Like RST, G&S analyze a text into an hierarchical structure. The basic idea of G&S is that a speaker or writer has a particular intention for uttering each constituent in the hierarchical representation of the discourse. The task of a discourse analyst is to recognize and to define these intentions, and to make decisions about dominance and satisfaction-precedence relations between them. Practical instructions for analyzing texts based on this theory are formulated by Nakatani, Grosz, Ahn and Hirschberg (1995). In our research three G&S experts were asked to analyze the same texts in terms of this practical guide. The analysts had access to the written transcripts only. The agreement between the three G&S experts was .73, .67, .68 and .68 for the texts respectively. These reliability results were significant too.

The main difference between RST analyses and G&S analyses is that the RST analyses are very detailed, generating very deep discourse structures, while the G&S analyses are less detailed, generating rather flat structures. With regard to prosody the question is whether the detailed RST annotations are reflected fully by the prosodic variation. If not, we might confine ourselves to more flat discourse analyses, like the one of G&S.

The second part of the research is about the correlation between the RST text structures of the four texts and their prosodic features. In a later stage we will use the G&S analyses to investigate their correlations with the prosodic features of the texts. We used one RST text structure per text to which the prosodic features were related. The spoken material was analyzed with a speech processing program. We measured pause durations at all segment boundaries and the maximum pitch of all segments (in Herz).

Our hypotheses about the relation between textual and prosodic features were as follows. (1) A higher segment boundary is associated with a longer pause. (2) A higher segment boundary is associated with a higher pitch maximum in the following segment.

There are several possible ways to test these hypotheses. For example, we may apply a relative method taking all adjacent segment boundaries in the text and look whether a boundary that is higher in the hierarchy is associated with a longer pause and a higher pitch maximum than a boundary that is lower in the hierarchy. Or we may apply a relative method taking segment boundaries that are adjacent in the hierarchy in terms of superordinate-subordinate relations and look whether a given (superordinate) boundary is associated with a longer pause and a higher pitch maximum than a subordinate boundary. Alternatively we may apply an absolute method and look for a direct relation between levels of text structure and prosodic features. We used these three methods.

According to the method of adjacent pairs, we inspected the prosodic features of all adjacent segment boundaries. On average higher segment boundaries are associated with a pause duration of .69 seconds and lower segment boundaries with a pause duration of .45 seconds (t(103) = -5.62, p < .001). In addition, segments following higher segment boundaries are associated with a mean maximum pitch of 215 Herz and segments following lower segment boundaries with a mean maximum pitch of 210 Herz (t(103) = -.72, n.s.). According to the other relative method, we inspected the prosodic characteristics of only those boundaries in the hierarchy that have a relation of subordination and superordination with each other. Based on 151 cases, Wilcoxon Signed Ranked tests confirmed both hypotheses (p < .001 for both pause duration and pitch maximum). According to the absolute method we correlated the levels of the segment boundaries in the hierarchies with the pause durations associated with these boundaries and the pitch maxima of the segments following these boundaries. The correlation between level in the hierarchy and pause duration was significant. However, pause duration showed a clear pattern for levels 1, 2 and 3 only. In other words, very long pauses occurred at level 1, shorter pauses occurred at level 2, and still shorter ones occurred at level 3. At lower levels in the hierarchical structure the pauses were

more or less equally short. The correlation between level in the hierarchy and pitch maximum was not significant. The behavior of the pitch maximum showed the expected pattern for levels 1 and 2 only. These findings confirm the hypothesis concerning the relation between text structure and pause duration. The relation between text structure and pitch maximum is significant as far as subordinate and superordinate pairs in the hierarchy are concerned. At the same time the results of the absolute method suggest that there is no direct mapping between the fine-grained structure as obtained with RST and prosodic structure.

References

Grosz B. & Sidner, C. (1986). Attentions, intentions, and the structure of discourse. Computational Linguistics, 12, 175-204.

Mann, A.C. & Thompson, S.A. (1988). Rhetorical Structure Theory: Toward a functional theory of text organization. Text, 8, 243-281.

Nakatani, C., Grosz, B., Ahn, D. & Hirschberg, J. (1995). Instructions for Annotating Discourses. Center for Research in Computing Technology, Harvard University, TR-21-95.

Sourcer's Apprentice: Facilitating Document-Supported History Instruction in the Classroom

J. van Dyke*, C. A. Perfetti* & M. A. Britt**

*University of Pittsburgh (U.S.A.)

**Northern Illinois University (U.S.A.)

jvandyke+@pitt.edu

History offers a literacy environment as rich as any a student is likely to encounter prior to college. The study of history centers on documents--letters, treaties, notes, official records, diaries--as well as textbooks. Instruction that makes good use of this rich text environment has the potential to support broad-based literacy skills that may extend beyond history classrooms to other cases of text learning, reasoning, and writing. It is not enough then merely to make documents available and to advocate their use. We must demonstrate, in standard classrooms, the value of document-supported history learning in a context that will promote actual student use of the documents. Both students and teachers must find document access to be useful and interesting. Toward that end, we have focused not just on developing an instructional tool that will encourage document use in the classroom, but also on the deployment of our tool in classroom situations where its effectiveness can be evaluated.

Sourcer's Apprentice is a Java application to promote an attitude of evidence seeking and evidence evaluation in students while developing a sense of document type and document privilege in history. It was designed to provide high school history students with opportunities to practice the kind of document-based reasoning exhibited by expert historians (Wineburg, 1991). Our goal was to create a simple coached-apprenticeship system (Gabrys, Weiner, & Lesgold, 1994) that would provide students the supports they need to interact with documents in a more authentic way. Although such systems are usually quite technically complicated (Lesgold, Lajoie, Bunzo, & Eggan, 1992), Sourcer's Apprentice is an example of how a coached apprenticeship system can be implemented in a very simple way.

This presentation will provide an on-line demonstration of the Sourcer's Apprentice, illustrating how the system can structure a student's inquiry into a historical problem. The Sourcer's Apprentice was built on a foundation of six principles which Cognitive Psychology has shown to be important for successful tutoring: Learn by problem solving, support expert representations, task decomposition, support transfer, provide explicit instruction, motivate engagement. We will demonstrate how each of these principles is realized in the Sourcer's Apprentice interface. In addition, a primary goal of our project has been not just to develop an interactive tool for history instruction, but to evaluate this tool in actual teaching situations. To date, the Sourcer's Apprentice has been used by lower-level "mainstream" students, intermediate level students described by their school as "average", and upper-level advanced placement students (Britt, Perfetti, Van Dyke & Gabrys, 1999).

We will present the results of a new study that examines how Sourcer's Apprentice should be integrated into the classroom. This study was designed so as to manipulate students' prior knowledge of the historical controversy to be explored. We found that prior knowledge did not affect the number of factual propositions students include in their essays, but it does affect the amount of original material in their essays. A corresponding increase in the number of connectives used by students with prior knowledge suggests that this new material represents more complex reasoning about the historical situation, including inferences about underlying causes and the motives for the actors involved. In addition, we have shown that even limited interaction with

Sourcer's Apprentice increases the number of propositions and the number of connectives students use in their essays. We conclude by suggesting that the Sourcer's Apprentice can be useful either as a way to augment pre-existing instruction or as a way to present new material in a novel and engaging format. In addition, this study implies that learning from texts is hierarchically structured. Basic knowledge of the situation and actors is important not only as a context for text comprehension, but must be acquired in order to facilitate complex reasoning about causal relationships.

Middle School Students' Processing of Multiple Accounts of an Historical Event

M. Wolfe, S. Goldman, C. Mayfield, P. Meyerson & D. Bloome

Vanderbilt University (U.S.A.)

Mike.Wolfe@Vanderbilt.edu

Research conducted over the past decade reveals that learners use a wide range of strategies to comprehend what they read. Think-aloud methods have revealed that adults and children who explain and elaborate what they are reading to themselves and who have a flexible approach to solving comprehension problems (i.e., use a variety of strategies) remember text and solve problems better than those who do not (e.g., Chi, deLeeuw, Chiu, & LaVancher, 1994; Coté, Goldman, & Saul 1998). Such activities reflect a "knowledge building" approach to creating a mental representation of the information conveyed in text. Successful knowledge builders engage in processing that transforms the text through self-explanations that rely on prior knowledge and active efforts to interconnect information from different parts of the text (Coté & Goldman, 1999). Knowledge building results in mental representations that are coherent, integrated with prior knowledge, and reflect the situation described by the text (Coté, & Goldman, 1999; Coté, et al., 1998). The vast majority of these studies have examined comprehension of individual texts.

However, in recent years, there has been growing interest in reading comprehension and learning across multiple texts (Bloome, et al., 1999; Goldman, et al., 1999; Perfetti, Rouet, & Britt, 1999). The multiple text focus is stimulated, in part, by research on disciplinary expertise that reveals that experts routinely engage multiple sources of information in the course of their work (e.g., Leinhardt & Young, 1996; Wineburg, 1994) whereas nonexperts do not. Experts integrate across documents, compare, contrast, and evaluate knowledge claims, and pay attention to the sociocultural context of the text (e.g., author, time, and place). Specifically in history, Wineburg (1994) reported that experts compare documents (corroborate), place them in their spatial and temporal contexts (contextualize), and consider the bias and intent of the authors (source). However, although university undergraduates noted the source of documents (Rouet, Britt, Mason & Perfetti, 1996), they tended not to create intertextual links among documents (Greene, 1994). Studies of student learning from social studies textbooks in primary and middle schools show that students typically learn lists of facts and acquire minimal understanding of causal principles and relationships among aspects of social systems (Beck & McKeown, 1994).

The present research examined the processing and knowledge building efforts of adolescents who read explicitly contradictory texts on the same historical event. Primary interest was in whether the patterns of processing and knowledge building were similar to those observed for the processing of single texts (Coté et al., 1998). Specifically would students process the texts as sequential, independent texts or would they draw connections between them and make reference to authors. In previous research we had found that students in this age range knew that a text reflected the perspective of the author but they did not use this kind of information in interpreting texts (Goldman et al., 1999).

Method

A total of 44, 12 and 13 year old students participated. They represented a range of reading comprehension skill levels. All students had been exposed to similar instructional units on ancient Rome within the three months prior to their participation. Students read two, short (approximately 200 words) explanations for the Fall of Rome, explicitly indicated as having been written by two

different historians. Each text explicitly stated disagreement with the other author's explanation, made a claim about their own explanation and provided evidence for their claim. Students provided think-aloud protocols as they processed each text. After reading, students generated their own explanations for the Fall of Rome. The think aloud protocols were parsed into event types and coded using a system similar to our previous scheme (Coté et al., 1998). The coding categories distinguish among paraphrases, monitoring statements, evaluations and various kinds of elaborative reasoning about the texts, including self-explanations and connections among text segments.

Results

The distribution of processing events for the sample as a whole indicated 23% paraphrases, 58% elaborations, 13% evaluations, 3% monitoring, and 3% problems. This is a higher rate of elaboration, a lower rate of monitoring, and about the same rate of paraphrasing and evaluating as we obtained with single text processing. Of the evaluations, 9% were of the author, indicating that students did not generally pay a great deal of attention to the author. Consistent with our previous work, about half of the elaborations were self-explanations that used prior knowledge. The other elaborations reinstated information from elsewhere in the text, from the first text that had been read, or prior knowledge that had been stated earlier in the protocol. Interestingly, 35% of the reinstatements were information from the first text that had been read, illustrating that students were connecting across the multiple sources. Furthermore, when students reinstated information from the previous or current text, 63% of the time they explained the connection. One interesting difference between the present and previous studies was in the percentage of prior knowledge elaborations that were categorized as irrelevant to the meaning of the text: 68% of prior knowledge elaborations were associations to personal experience. In past work, the topics of the texts made it more likely that associations to personal experiences would be relevant to the text interpretation. For historical accounts, personal experiences were less likely to be germane to the authors' knowledge claims. The findings for the group as a whole mask some of the interesting differences in knowledge building approaches taken by individual and clusters of students. In ongoing analyses we are looking at processing profiles to better understand these differences.

Conclusion

Adolescents generally approached the processing of multiple historical explanations for the fall of Rome in a manner similar to single text comprehension. The differences, however, suggest that students were engaged in some critical evaluation of the claims being made in each text. Students did little reasoning about the authors of the texts, but did make attempts to integrate the content of the information across the two texts. These comparisons provide a basis for developing students' skills in comprehending, evaluating and subsequently constructing evidence-based accounts of historical events.

References

Beck, I. L., & McKeown, M. G. (1994). Outcomes of history instruction: Paste-up accounts. In M. Carretero & J. F. Voss (Eds.), Cognitive and Instructional Processes in History and the Social Sciences (pp. 237-256). Hillsdale, NJ: Erlbaum.

Bloome, D., Goldman, S. R., Meyerson, P., Coté, N., Wolfe, M. B. W., & Mayfield, C. (1999, April). Reconceptualizing Reading as Intertextual Practice. Paper presented at the meeting of the American Educational Research Association, Montreal.

Chi, M. T. H, deLeeuw, N., Chiu, M., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. Cognitive Science, 18, 439-477.

Coté, N., & Goldman, S. R. (1999). Building representations of informational text: Evidence from children's think-aloud protocols. In H. Van Oostendorp & S. R. Goldman (Eds.), The Construction of mental representations during reading (pp. 169-193). Mahwah, NJ: Erlbaum.

Coté, N., Goldman, S. R., & Saul, E.U. (1998). Students making sense of informational text: Relations between processing and representation. Discourse Processes, 25, 1-53.

Goldman, S. R., Meyerson, P., Wolfe, M. B. W., Mayfield, C., Coté, N., & Bloome, D. (1999, April). "If it says so in the text book, it must be true:" Multiple Sources in the Middle School Social Studies Classroom.Paper presented at the meeting of the American Educational Research Association, Montreal.

Greene, S. (1994). Students as authors in the study of history. In G. Leinhardt, I. Beck, & K. Stainton (Eds.) Teaching and learning history (pp. 133-168). Hillsdale, NJ: Lawrence Erlbaum.

Leinhardt, G., & Young, K. M. (1996). Two texts, three readers: Distance and expertise in reading history. Cognition and Instruction, 14 (4), 441-486.

Perfetti, C. A., Rouet, J-R., & Britt, M. A. (1999). Toward a theory of document representation. In H. Van Oostendorp & S. R. Goldman (Eds.), The Construction of Mental Representations During Reading. Mahwah, NJ: Erlbaum.

Rouet, J. F., Britt, M. A., Mason, R. A., & Perfetti, C. A. (1996). Using multiple sources of evidence to reason about history. Journal of Educational Psychology, 88 (3), 478-493.

Wineburg, S. S. (1994). The cognitive representation of historical texts. In G. Leinhardt, I. L. Beck, & C. Stainton (Eds.), Teaching and learning in history (pp. 85-135). Hillsdale, NJ: Lawrence Erlbaum Associates.

Using Shared Physical Space to Ground Analogical Models

R. A. Engle

Stanford University (U.S.A.)

randi@csli.stanford.edu

When I asked undergraduates to explain to a partner how deadbolt locks work, they spontaneously spoke while gesturing, drawing diagrams, or demonstrating with a lock. In this paper, I claim that these multimodal explanations supported the construction of multiple analogical models of locks, each one associated with distinct regions within the physical space shared by interactants. Almost all the gestures, diagrams, and lock demonstrations contributing to a given model were located in the same region of shared space. In contrast, visible signals contributing to different models were usually performed in different regions. Systematically locating each visible signal in a particular region of shared space may have helped participants collaboratively ground the explanations (Clark & Schaefer, 1989) by helping addressees identify which model each multimodal utterance was updating.

Method

Six English-speaking undergraduates examined a sample lock and studied a written explanation of how locks work. Each undergraduate then explained locks to a series of three individual undergraduates who did not know how they work. For one explanation, the sample lock was available; for another, paper and pen were; for a third, no props were available. Order of prop availability was counterbalanced. Having different props available was included to mimic the variation in real-world explanations. All 18 explanations were videotaped by three cameras, two horizontal on each participant, and one vertical focused down on the table where participants gestured, drew diagrams, and manipulated the lock.

Analysis

Transcription and sampling: Speech was coded for pauses, false starts, and intonational contours. All hand movements including gestures, manipulations of the lock, and drawing or writing actions were described and then timed to the nearest syllable of ongoing speech. One episode from each explanation was intensively analyzed. Each episode was a coherent piece of the explanation that had been collaboratively grounded as a "conversational contribution" (Clark & Schaefer 1989). On average episodes contained 7.1 intonation units and 6.2 visible signals, a total sample of 127 intonation units and 111 visible signals.

Coding models: Each multimodal utterance (i.e., a visible signal and its associated intonation unit) was coded as updating a particular analogical model. This was done sequentially starting at the beginning of each episode. For each utterance, I determined whether any of the existing models could be updated with the information from the utterance while still forming a coherent model. To be coherent, a model had to: (a) include a connected set of objects, (b) use a single spatial perspective (e.g., side or top view, but not both), (c) be relevant to no more than one continuous time span (e.g., the lock's layout before or after the key is inserted, but not both), and (d) be specified at the same level of abstraction. An utterance was considered to establish a new model if

a coherent model that satisfied all 4 conditions could not be formed by updating any existing model.

Identifying spatial locations: Using horizontal and vertical video displays, the outer boundaries of each visible signal were traced onto transparencies. Transparencies for different visible signals were superimposed over each other to determine whether pairs of visible signals were located in the same region or a different one. For each visible signal it was determined: 1) whether it was located in the same region as previous visible signals contributing the same model, and 2) whether it was performed in a different region from previous visible signals used to construct different models. To count as a different region, the locations could not intersect; to count as the same region they had to.

Results

Using the same region for each model: Did explainers consistently use the same region to locate each visible signal that updated the same model? They did. Of the 87 visible signals updating an already existing model, 94% were located in a region that intersected the region used by previous visible signals contributing to that model. In fact, the locations of most visible signals updating the same model almost completely overlapped each other. There was no difference between conditions in how often the same region was used for the same model's visible signals ($F(2, 84) = 2.13$, $p = .13$)

Using different regions for different models: If models are located in shared space, then visible signals contributing to different models should be located in different regions of the space. In the sample, there were 132 pairs of visible signals that contributed to different models. Of these, an average of 81% were performed in completely non-intersecting regions. (Most exceptions only intersected slightly, suggesting that explainers at least made sure that different models were centered around distinct loci.)

More consistently using different regions in the paper and lock conditions: Visible signals for different models were more likely to be kept in non-intersecting regions in the lock (95%) and paper (97%) conditions than in the no props condition (63%, $F(2, 129) = 13.95$, $p < .0001$). Having the lock or paper provided new places for locating models which included permanent artifacts useful for keeping track of model locations. Without the lock or paper in the no props condition, participants could only locate their models with gestures, which also generally leave no trace after being performed.

Discussion

In a large majority of cases, the spatial location in which explainers performed a visible signal was informative about which model was being updated by that visible signal and its associated speech. Different models were almost always located in different regions of lock, paper, and gesture space. These results extend previous findings on the use of space in gesture (McNeill & Levy, 1993) to include more kinds of visible signals (i.e., also diagrams and demonstrations) that establish more complex referents with speech (updates to analogical models). This use of space may also be similar to ASL and other signed language practices of distinguishing referents using spatial location (Emmorey, Corina & Bellugi, 1995; Liddell, 1995). Systematically locating visible signals in shared space may help interactants keep track of which model is being updated so they can coordinate their interpretations. In future work, I plan to test this hypothesis by seeing whether addressees' choices about which model to update are indeed sensitive to where speakers locate their visible

signals. If they are, then locating visible signals in particular spatial locations is another mechanism people use to establish shared understandings in discourse.

References

Clark, H.H., & Schaefer, E.F. (1989). Contributing to discourse. Cognitive Science, 13, 259-294.

Emmorey, K., Corina, D. & Bellugi (1995). Differential processing of topographic and referential functions of space. In K. Emmorey & J.S. Reilly (Eds.), Language, gesture, and space (pp. 43-62). Hillsdale, NJ: Erlbaum.

Liddell, S. (1995). Real, surrogate, and token space: Grammatical consequences in ASL. In K. Emmorey & J.S. Reilly (Eds.), Language, gesture, and space (pp. 19-41). Hillsdale, NJ: Erlbaum.

McNeill, D. & Levy, E.T. (1993). Cohesion and gesture. Discourse Processes, 16, 363-386.

Characterizing Discourse Modes with Linguistic Tools

C. S. Smith

University of Texas (U.S.A.)

carlotasmith@mail.utexas.edu

I will discuss local structure in written texts, taking a linguistic approach. I argue that we should recognize a level of structure, the discourse mode, which is intermediate between genre units, the units of discourse coherence relations, and sentences. Discourse modes are realized by stretches of text with characteristic semantic and pragmatic meanings. The meanings are licensed by composites of linguistic forms.

On encountering a discourse, one uses a great deal of world knowledge to understand it. Each discourse represents an activity of a certain kind with its own function and structure (Levinson 1979). But genre is not sufficient to account for what one knows about text structure. There are stretches of text which are intuitively of different types, e.g. Narrative, Description, Commentary, Argument. They tend to have a particular force which is due to cluster of linguistic forms and their interpretation. Stretches of text with different clusters realize different discourse modes. The discourse modes cut across genre lines. I will show that discourse modes can be identified through close analysis of a text. Linguistic forms give rise to semantic and pragmatic interpretation. I will use the term 'interpreted linguistic features' to include linguistic forms and the interpretations which they trigger, or license. There are key interpreted linguistic features which characterize each mode. I will investigate the five modes of Narrative, Description, Report, Expository-informative, Argument-commentary. In order to ensure a relatively small set I will allow for considerable variation within a mode. The modes vary in point of view, or authorial stance, and in register, or level of formality.

The notion of discourse mode accounts for the fact that actual discourses are rarely monolithic: They tend to contain passages of different kinds. Narrative sequences, for instance, appear in fiction, history, newspaper editorials, letters, etc. And in a narrative fictional text one finds descriptive passages, and perhaps commentary, as well as event sequences. Narrative, description and commentary make different contributions to a text; they also have different linguistic features. The discourse modes can be characterized according to three key interpreted linguistic features: (i) the types of entities they introduce; (ii) temporality; and (iii) principles of local organization.

(i) Types of entities. Type of entity is intended in the sense of Discourse Representation Theory: The sentences of a text licenses the introduction of various entities - including individuals, events, and times - into the developing semantic structure that represents the meaning of the text (Kamp & Reyle 1993). The entities that figure in this analysis are situations in the world, or events and states; and abstract entities such as complex statives, facts, propositions and perhaps others (depending on level of detail). The linguistic forms that license recognition of these entities are complex, usually involving a verb and its arguments. Each type of entity constitutes a covert linguistic category with unique properties, distributional and semantic (Vendler 1967; Smith 1991; Asher 1993; Peterson 1997). Very generally, texts in the modes of Narrative and Description primarily involve situations; Report and Exposition primarily involve situations and facts; Argumentation has facts, propositions, and often situations.

(ii) The temporality of a text depends to some extent on the entities it introduces. Events and stage-level states are located in time, whereas abstract entities (individual-level states, facts, propositions) are atemporal. If located in time, situations are static or dynamic. Temporally located entities appear mainly in the Narrative and Report modes, both dynamic; and Description, which is

static. The atemporal discourse modes are Exposition and Argumentation. Relevant linguistic forms are tense, adverbials, constellations of verbs and their arguments (Smith 1980, 1991, 1998).

(iii) The third interpreted feature is the principle of local organization. Principles differ among the modes. Narrative is organized in terms of narrative advancement, which is fairly well understood. Narrative events are sequentially ordered; narrative time advances with each successive closed event, and does not advance otherwise. Thus narrative is organized in terms of temporal location, assuming the pervasive analogy between space and time. States involve metaphorical location and motion, focusing on what is predicated of the primary figure and how the components of the situation pertain to the figure (following Talmy, 1985). This idea provides a basis for understanding of certain non-narrative modes.

Texts in the Descriptive mode advance from one part of the scene to another; more globally, from scene to scene. Texts in the Report mode are organized either sequentially according to the reporter's perspective, by changes of time or space, or in terms of the domain talked about. Reports tend not to present events sequentially but are organized by the reporter, at Speech Time (Caenepeel, 1995).

Extending the approach to abstract entities, I assume metaphorical location and motion within a domain. The principle of advancement for modes with abstract entities involves metaphorical change of location, motion from one part of a given domain to another. In the Expository-Informative mode, the domain is often organized by the material talked about; in the Argument mode, the form of the argument usually organizes the material. Thus a stretch of text in this mode may move from claim to evidence or, more globally, from claim to claim.

Metaphorical movement from one part of a domain to another is conveyed by linguistic cues of different types. For texts with abstract entities the linguistic forms that realize dynamism of narrative time are not available. Cues include shifts from one type of entity to another, e.g. from Fact or Proposition to Event. Other cues involve information structure: change of sentence topic is an important cue. Pragmatic inference is often required to understand such shifts. Position is only one of several kinds of cues to sentence topic in English texts. Topic in English is determined by a confluence of factors, including the structure of adjacent sentences. Cues include word order, lexical repetition, syntactic structures that defocus, and patterning with other sentences. The points are illustrated with discussion of real examples.

I assume that receivers assemble and interpret the different cues in a text, e.g. lexical choices, syntactic and information structure, patterns within and across sentences. Pragmatic inference is used frequently as one arrives at an interpretation. The preferred interpretation is most compatible with all the information and relevant pragmatic knowledge. I do not attempt to state the processes involved.

References

Caenepeel, M. (1995). Aspect and Text Structure. Linguistics, 33, 213-253.

Kamp, H., & Reyle, U. (1993). From Discourse to Logic. Dordrecht: Kluwer.

Peterson, P. (1997). Fact, Proposition, Event. Dordrecht: Kluwer.

Smith, C. S. (1980). Temporal Structures in Discourse. In N. Rohrer, (ed), Time, Tense, and Quantifiers.

Smith, C. S. (1991). The Parameter of Aspect. Dordrecht: Kluwer.

The Scientific Status of Rhetorical Structure Theory: Two Views

W. C. Mann* & S. A. Thompson**

*SIL International (U.S.A.)

**University of California Santa Barbara (U.S.A.)

billmann@sprintmail.com

In the 1980s, Rhetorical Structure Theory (RST) was developed as a way to explore the structure of written monologues. Typically an RST analysis starts by dividing a text into some minimal units of interest, such as independent clauses or orthographic paragraphs. It assigns a role in the text to each such unit, primarily by linking parts of the text together using relations, and aggregating related parts into spans. Spans can be linked to other units or spans, and so that the text is connected together into a hierarchic structure. Most of the relations are asymmetric, linking a nucleus to a satellite. The largest span created in this manner encompasses the whole text. One use of RST has been as an account for the coherence of written texts. Coherence of a text can be seen as the text having an evident role for every part. Negatively, it is the absence of non sequiturs. RST analyses are created by people (observers) who make a structured set of judgments about a written text. An RST diagram is equivalent to affirmation of a particular set of judgments, which can be identified without consulting the observer. Thus each analysis has a definite interpretation. RST has been used more or less directly in a wide range of work. In addition to studies involving coherence, it has been used to study inter-clausal relations, conjunctions, clause combining, implicit communication, writing style and genre. In computation it has been used in text generation, summarization, text indexing and modeling of comprehension. The scientific status of RST has been found to be unclear, which has led to confusion and inappropriate expectations. A better understanding of this point would clarify notions of success, progress, appropriate rigor, completeness of accounts, degree of demonstration, prediction, confirmation, significance. This paper is intended to contribute to the understanding of these matters.

The paper presents two views of RST. In one view, RST is simply a way of identifying and making precise certain phenomena of language use. It collects data, and the data collection activity encounters patterns which help us to understand how language functions. It is a partially subjective process, involving reading and producing a diagram of text structure. RST in this view does not make claims or build hypotheses about how language works. We will call this the data gathering view. In the second view, RST is a partial model of how people structure written texts, or how they seek structure in written texts. Such a partial model could in principle be combined with a collection of other partial models to form theories of language use. Such theories would be subject to the normal scientific processes of formulation, test, confirmation, refutation and revision. We will call this the theory component view.

RST was defined with principal concern for the data gathering view. The developers were particularly concerned that each RST analysis has a definite interpretation. This was carried out partly by using definitions for relations, so that the process of text analysis is essentially application and affirmation of definitions. In order to make RST analysis more reliable, and also to increase the tendency toward agreement between observers, requiring observers to make difficult judgments has been avoided. As a result, the relations and structures of RST are more irregular than they otherwise would be. For the data gathering view, we examine the scientific interpretation of some of the patterns that have been identified during extensive RST analysis of texts. These include the very high proportion of texts that have analyses, tendencies for analysts to agree, low perceived ambiguity of discourse structures, relationships between coherence of a text and the existence of an analyses, the wide diversity of text genres for which RST is effective, patterns of

explicit and implicit indication of discourse structure, and ways that authors communicate implicitly using the discourse structure. These are not simply issues of methodology. Rather, the fact that the analysis process works tells something about what it is working on, i.e. about the nature of text. For RST as data gathering, it should be as "theory neutral" as possible, but complete neutrality is not possible. There are assumptions built into any text analysis method. A collection of assumptions inherent to the data gathering view are examined, including assumptions about language, text, observers and authors. The other view, the theory component view, takes a different approach to text structure.

Text structure has been seen in many ways, sometimes in isolation and sometimes as part of a larger linguistic scope such as semantic structure, pragmatic structure or communication. The underlying abstract nature of text structure -- its theory, has been considered by some to be a well-formedness notion, akin to a text grammar. Another suggestion is to see text structure as aggregations of cohesive links. Compositions of speech acts have also been suggested, as well as expression of particular genre organizations, the semantics of discourse markers, or an outworking of necessary rhetorical features. RST does not follow any of these suggestions. Another group of suggestions, which are closer to RST's methods, says that readers or hearers seek to perceive the goals, intentions or desired effects of the producers of text, that text structure expresses some of those intentions, and that writers rely on this intention-seeking activity of readers.

The theory component view of RST is probably most easily conceived as a small portion of a model of text creation. This follows the first major use of RST in computational linguistics, which was to organize computer-based text generation. Partly because of this use, RST has sometimes been expected to be more regular, objective and formal than it really is as a data gathering approach. RST cannot be a whole model of text creation, of course, since it does not deal with the roles of words, grammar or semantics, and it does not touch many parts of pragmatics. For these reasons, RST cannot function in isolation in this role. In an overall model of text creation, RST will be accompanied by other components. Furthermore, RST does not specify how it would be used in text creation. It merely suggests guidelines for ways that coherence might be achieved and how intentions of the author could partially shape text construction activity. Because RST does not specify how it would interact with other formative forces, it needs to be extended and reconciled to other bodies of knowledge. When such extensions are made, there is naturally a need within RST for regularization and augmentation to meet the demands of the framework adopted for the model as a whole. Some of the projects that have used RST have clearly seen a need for this sort of development.

In summary, scientifically, RST is a way of producing particularly examinable cases of text affecting people, yielding phenomena and structured data rather than a model of how such cases arise. The work of analyzing texts produces insights and surprises about text structure and function, enriching the phenomena to be accounted for. RST contributes ideas about what might be included in theories (or models or explanations) of language creation and reception, but it is at most an incomplete fragment of such a theory.

Emotional Inferences, Course of Activation in the Landscape Model

S. Groen, A. Syssau, F. De la Haye & D. Brouillet

University of Paul Valéry (France)

groen@smr1.univ-montp3.fr

Since the 70's, research on inferences in text comprehension aroused great interest, and this, more particularly, since the experiments of Bransford and Franks (1972) who showed that in addition to information elements present in the text, elements of the readers, knowledge are activated during reading. Indeed, successful reading comprehension involves not only the representation of the textbase in memory, but at the same time, the activation of readers' background knowledge, i.e. implicit elements that are inferred in order to obtain a coherent representation of the text. These inferences are a part of the relatively stable cognitive representation of what the text is about: The situational model (van Dijk & Kintsch, 1983, p. 337).

In our work, we are interested in the activation and maintenance of activation of knowledge about human emotions and the use of the activated knowledge to build mental representations of fictional characters' emotional states. If research on inferences are numerous today, the particular type of emotional inferences received only little study. These inferences correspond to the clarification of the emotions felt by the protagonists in response to the events or actions reported by the text (Graesser, Singer, & Trabasso, 1994). Three characteristics of emotional inferences seem to emerge from literature today. Emotional inferences would be generated during reading (on-line) and not once the reading accomplished (off-line). Indeed, in a succession of three experiments, Gernsbacher, Goldsmith and Robertson (1992) showed that the reader activates his knowledge about human emotions when reading a narrative, in order to infer the protagonist's emotional state. In extension to Gernsbacher et al.'s (1992) research, de Vega, Léon and Diaz (1996) specify that, not only emotional inferences are produced during reading, but they remain activated in working memory throughout the reading of the text. Emotional inferences would thus make it possible to ensure establishment of coherence when the text presents ambiguities or inconsistencies such as the introduction of new characters, changes of places or temporal discontinuities. De Vega et al., (1996) postulate that the interest that the reader carries to characters' emotional states supports integration of different portions of the text. From this point of view, emotional inferences would contribute to the development of a global coherent text representation. Results of another experiment of de Vega et al., (1996), permits to advance that emotional inferences are dynamic, that is, once produced, they are updated after a change in the protagonist's emotional situation in the continuation of the story.

We have examined the course of emotional inferences activation, within the framework Landscape Model (van den Broek, Risden, Fletcher & Thurlow, 1996). Indeed, this model of text comprehension permits us to study the inferential processes that take place during reading in considering inferences, activation in terms of a continuum of activation, and the way in which these processes lead to the construction of a relatively stable memory representation of a text. So we have been able to explore at the same time, the moment of emotional inferences' production and the maintenance of their activation in memory. A conventional way of representing a successful comprehended text in memory is in terms of a coherent structure. This structure can be represented as a network of nodes (corresponding to concepts) that are linked by anaphoric and causal connections (Trabasso & Suh, 1993; van den Broek & Lorch, 1993). On the force and the number of connections associated with a particular concept depend the probability of its recall (Graesser & Clark, 1985; Miller & Kintsch, 1980; Trabasso et al., 1984; Trabasso & Sperry, 1985). Moreover, considerable research has demonstrated that concepts that are activated simultaneously are connected in memory representation (Fletcher & Bloom, 1988; Fletcher, Chrysler, van den

Broek, Deaton, & Bloom, 1995; Kintsch & van Dijk, 1978). Thus, activation and co-activation of the different concepts conveyed by a text, can be determining in the recall of a text.

In considering the factors which determine the recall of concepts of a text according to van den Broek and al., on the one hand, de Vega and al.'s considerations (i.e. emotional inferences, activation maintenance throughout the text), on the other hand, we can advance that emotional inferences will have a better score of recall than do more neutral inferences. Indeed, emotional inferences are maintained activated longer, which according to authors of the Landscape Model would generate an important rate of total activation and, consequently, a high rate of recall. Moreover, if emotional inferences are maintained activated in memory during a more significant number of cycles, they will be probably strongly connected to other concepts, increasing their probability of recall. In sum, we expect that emotional inferences occupy an important place in the representation of the text in memory. We proceeded in two stages. First, we have employed the procedure developed by the authors of the Landscape Model. That is, we have submitted four emotional and four neutral stories to two groups of ten participants that noted activation scores for each word and several inferences on an activation matrix. The first group of participants (called "expert") received preliminary training of the contents and the application of the activation rules, the second group (called "naive") was not informed of these rules. Results show a) that the Landscape Model efficiently predicts activation of concepts during reading (i.e, results of the two groups correlated at r = .94), and b) emotional inferences are maintained activated longer than predicted the model.

Second, we proceeded to the recall tasks. We solicited 120 volunteer psychology students to whom we submitted the four emotional and four neutral stories (each participant reads one story). In order to measure the memorizing of the texts, we replicated the step of van den Broek et al. We submitted the narrative texts by giving the following instruction: " Read this story very attentively during one minute ". Immediately after reading the text, we gave the participants this second instruction: " Mark on a sheet of paper, all you remember of the previously presented story ". Then, after a one hour's delay, the same instruction was given. And finally, a supplementary task of recall was effectuated after a week's delay. We also proceeded at a recognition task at that moment. We presented six sentences to the participants: One sentence containing the emotional and neutral inferences, one sentence that made part of the story, two sentences that did not made part of the story, and two paraphrases. Participants had to answer if yes or no these sentences made part of the story they read a week before. And, in addition they had to indicate on a five point scale how sure they were of their answer. By effectuating this recognition task we were able to distinguish the accessibility and availability of our inferences. Results indicate a general superiority of emotional inferences recall in comparison with more neutral inferences. On the contrary, there is no difference for the two types of inferences concerning the recognition task.

The Influence of Distance and Sufficiency on the Production and Maintain of Forward Inferences

M-P. Quintana*, I. Tapiero* & P. van den Broek**

*University of Lyon 2 (France)

**University of Minnesota (U.S.A.)

marie-pilar.quintana@univ-lyon2.fr

This research is part of the studies investigating the processes and conditional production of causal inferences during reading, leading to a causal network representation of the text: Namely the ones on the Causal Inference Maker model (van den Broek, 1990; van den Broek, Fletcher & Risden, 1993). In this model the causal strength of a relation and the distance between two statements in the text surface structure determine the type of causal inference produced during reading, mainly forward and backward inferences. Our purpose was to deeper investigate the influence of sufficiency and distance between two statements on the production of forward causal inferences during narratives reading. Indeed, van den Broek and his collegues showed that the more sufficient a relation is, the more likely it is for a predictive forward inference to be made, which means activation of a highly probable consequence. If the reading of a cause is due to lower sufficient relation but with high necessity, the reader can infer the importance of the event and maintain it activated until its consequence appears, which is another type of forward inference. Therefore the cause and/or the consequence can be maintained activated during the production of forward inferences. The questions that arise are (1) the possibility for a forward inference to become backward under the influence of the distance between the cause and the effective appearance of the consequence in the text surface structure. Is it a decay of the cause activation or of the probable consequence one? (2) What are the events activated and maintained during the production of predictions: The cause, the consequence, or both? Varying the distance between the cause and the appearance of its consequence in the text surface structure, and varying the sufficiency of those relations should allow us to determine the conditions under which the cause and/or the consequence are activated and maintained.

Method

To investigate this theoretical questioning, we used an on-line probing task on words refering either to the cause event or to the consequence event. The reading was interrupt just before the reading of the consequence. Two causal relations were tested for each narrative (two interruptions per narrative). The causal relations could be adjacent (one sentence), semi-adjacent (five sentences) or non-adjacent (ten sentences) in the text surface structure. Moreover they could satisfy to high sufficiency (physical causality) or to medium sufficiency (psychological causality). Necessity was always satisfied. Therefore we constructed 12 simple narratives of 22 sentences in 3 versions so as to have the same statements in the three distances. Then, these three versions of narratives were transformed so as to present less expected consequences, but satisfying to the same type of causal relation. We had then 6 versions of the 12 narratives. To select the most " fitting-to-naïve readers " consequences, we did a pre-test in which we asked naïve readers to complete small texts presented, with either a plausible consequence or a less plausible consequence but always true according to what they read before. As expected the answers for the physical causality relations were mostly the same. Differences appeared for the psychological relations, as sufficiency is lower, several probable consequences are possible. We used in our material the most often given concepts. Then for the probing task, we selected words refering to either the cause or the consequence, created non-words, and selected other words from the french language. A participant tested on cause words for a given relation, could not be tested on consequence words for that same relation. Therefore at each interruption, a participant had to judge whether the randomly presented words were words or non-words. The reading times and the reaction times of the probe task were recorded. Our main predictions on reaction times were that: (i) For consequence words refering to high sufficient statements, we expected no difference between the words cause and consequence, i.e., forward prediction. (ii) For medium sufficiency, we expected reaction times on consequence words to be lower than for cause words. And this difference would increase with distance. This decay of activation should be the result of the change of type of causal inference, i.e., from forward to backward. The distance semi-adjacent being the central point of activation decay.

Results

Our results confirmed our main hypothesis. Probe words were rated faster than words. We found no significant effect of the interaction Cause * Consequence. But within reaction times on probe consequence words, high sufficient causal relations showed that consequences are maintained and integrated in memory as the distance increased. Reaction times decreased (but not significatively) according to the increase of distance. Therefore, forward inferences highly sufficient are maintained and strengthened through distance (assumption i). High sufficiency leads to the production of purely forward inferences (predictions). For medium sufficient causal relations, we found an increase in reaction times on probe consequence words under the influence of distance. Therefore, consequences activated are progressively deactivated as the reader proceeds through the text without encountering the effective consequence. A forward inference can become backward under the influence of distance (assumption ii). The distance semi-adjacent appeared to be the central point of activation decay/integration since we did not find any difference in reaction times for medium and highly sufficient consequences activation. Within reaction times on probe words cause, results showed a different pattern of activation for cause statements. It seems that probe words refering to highly sufficient causes are rated faster than the ones refering to medium sufficiency statements. Both strengths showed the same pattern of activation according to distance, i.e., progressive deactivation as the distance increases (assumption ii). With highly sufficient causal relations, a forward predictive inference was made and maintained regardless of distance (assumption i).

Discussion

As a conclusion, we can assume that forward inferences consist in maintaining both cause and consequence events, i.e., the relation between the two events. Readers then, used their causal knowledge to predict future events. As long as they proceed through the text, the predicted relation is more deeply integrated in the network under construction, therefore when an effective unexpected consequence appeared in the text, they had to create a new connection, which takes more time. The lower the sufficiency is, the more different probable consequences can be infered. Nevertheless, our results on reading times, confirmed that the activation of the consequence is less important in forward inferences with lower sufficient relations. Cause and consequence being equally activated in the production of predictions with (notably) highly sufficient relations, we can expect that in some particular cases, a symetrical connection could be represented in memory. The same causal relations could be used in a priming task, according to the three distances, and presented in a related and reversed way, so as to investigate this possibility.

Reference

van den Broek, P. W. (1990). The causal inference maker: Towards a process model of inference generation in text comprehension. In D.A. Balota, G.B. Flores d'Arcais, & K. Rayner (Eds.), Comprehension processes in reading (pp.423-446). Hillsdale, NJ: Erlbaum.

Causal Inferences in the Comprehension of Scientific Text: The Role of Causal Connectives

P. Maury & A. Teisserenc

University of Paul Valéry (France)

pmaury@smr1.univ-montp3.fr

The expression of a causal relationship in a text can either be implicit and suggested by juxtaposed clauses or on the opposite, explicitly signaled by causal connectives such as because or and. Several studies using sentence pairs have shown that the presence of an interclausal connective facilitated the construction of a coherent local relationship between the two sentences. For Millis &Just (1994), the causal connective because triggered the reactivation of the information read in the first clause and allowed the integration of the two clauses in a single representation. However, during the reading of a real text, several lines long, the readers normally attempt to achieve coherence at both local and global levels. Therefore, it is sometimes necessary when reading a text, to connect the current sentence to previous knowledge in long term memory as in this exemple from McNamara et al., (1996): " The blood cannot get rid of enough carbon dioxyd through the lungs. Therefore, the blood becomes purplish. " In spite of the presence of the connective therefore, the readers are not able to construct a coherent representation for these two sentences because they failed to integrate the information currently read to background knowledge as for exemple " the blood that flows from the heart to the lungs is low in oxygen and looks blue whereas the blood that flows out of the lungs is high in oxygen and looks bright red " (McNamara et al., 1996, p. 5). Thus, the role of connectives on scientifc texts appears to be quite specific because in narratives, it is easier for a reader to infer interclausal relationships from the content of the texts whereas in popular scientific texts, the construction of these relations is only based on cohesion cues and not on pragmatic coherence (Mac Cutchen, 1987). Millis, Golding and Barker (1995) described a facilitative effect for causal inference generation during the reading of pairs of narrative statements connected with because. However, one can argue that the sentences to be read included a coherence break that has possibly contributed to the inference generation. So, the aim of this study is twofold : We determined first whether the presence of a causal connective increase inference generation in scientific texts too; then, we raised the question of the durability of this activation effect by comparing the results of an on-line completion test and an unexpected cued recall test. In other words, we wondered whether causal inferences activated during reading are encoded and integrated in the representation of what the text is about.

Method

Subjects. A total of 135 undergraduate subjects took part to this experiment. They were split in four experimental conditions "because condition","therefore condition", "and condition" and "a constrained word stem completion condition (without reading any text)"

Material. Nine experimental popular scientific texts and 6 filler texts were displayed. Attention was payed that none of the filler texts described similar technological or natural mechanisms as those evocated in the experimental texts. At the beginning of each text, there is no title and no superordinate goal explicitly mentionned in the introductory sentences. Specific background knowledge are not necessary to understand the texts because they rely on basic principles of physics or chemistry learned at secondary school. The target word occurred after the fourth sentence in the experimental texts and after the first or the second sentence in the filler texts.

Procedure. Subjects were instructed to read carefully the texts in order to answer final comprehension questions (not relevant to the details of each story). The sentences were either conjoined by a connective (therefore, and, because) or separated by a period. Subjects read the texts at their own pace in a self paced fashion. After the last segment of the sentence had been presented, the string "XXX" appeared on the screen and then replaced by the first two letters of a

target verb. Subjects were instructed to give orally the first verb that came to their mind and fit the first capital letters of the target verb. Then, after the whole reading, subjects were asked to answer comprehension questions on the filler texts as soon as possible. Finally, an unexpected final cued recall test was proposed to the subjects in the three experimental conditions. They were instructed to write down the sentence in the text that immediately followed versus preceded (for half the subjects of each experimental condition) the clause displayed on the center of the screen. They were told that if they could not remember the exact words, they should write down any words or ideas that they could remember for the first versus second sentence.

Results and Discussion

We developed here only the relevant results to the verb completion test and the recall data. Verb completion scores were submitted to a mixed analysis of variance with one between subjects factor (the experimental condition: "because", "and", "therefore" and "a constrained word stem completion condition ") and one within subjects factor (with or without connective). The results showed an effect of the experimental condition [$F(3, 104) = 5.88$, $p = .01$] with the highest score of completion in the "because condition" and the lowest in the "constrained word stem completion condition". The connective "therefore" lead to a score of completion significantly different from the "because condition" [$F(1, 104) = 6.62$, $p = .01$]. However, the difference between the "and" condition" and the "because condition" did not reach significance [$F(1, 104) = 2.65$, NS]. This finding suggests that causal inferences are activated during reading. The with or without connective effect is not significant [$F(2, 85) = 2.51$, NS] although the verb completion scores are higher in the without connective condition obtained in the three experimental conditions. In addition, the interaction did not reach significance [$p > 1$]. This result fits the connective interference-based interpretation proposed by Millis, Graesser and Haberlandt (1993) about connectives and memory for text.

Recall data: To be categorized as inferential recall, the subject had to write the inferential verb or expression corresponding to the cue rather than the sentence that was actually read.

Table 1. Mean inferential recall as a function of the three experimental conditions.

Experimental Conditions	Recall of clause 1	Recall of clause 2
"And condition"	3.16	4.75
"Therefore condition"	2.5	3.98
"Because Condition"	0.88	3.21

For exemple, Table 1 indicates the mean number of inferences that the subjects recalled. It is worth to notice that in the "because condition", the first clause refers to the consequence of the action while the second clause refers to the cause. The opposite pattern happens in the "and & therefore conditions".

The results showed that inferential recall score is highest for the second clause whatever the experimental condition. We observed that in the "because condition", the subjects failed to recall the inferential consequence of the action. Relationship between Comprehension and Memory for

Text A very few number of inferences activated during reading are recalled in the unexpected cued recall test. However, we noticed that a great proportion of inferences mentioned in the recall test were not activated during reading. This finding shoud suggest that causal inferences are briefly activated during reading rather than integrated in the representation of what the text is about.

Activation and Integration Processes Involved in Predictive Inferences Generation

N. Campion & J. P. Rossi

University of Paris XI (France)

Campion@limsi.fr

The present study examined the relevance of the distinction between two processes involved in the generation of predictive inferences. First, the activation processes which activate meaning that are associated to text information in readers semantic memory. Second, the integration processes by which the activated information becomes a component of readers' episodic text representation in long term memory. That distinction is central to the Construction-Integration Model (Kintsch, 1988) and to the Landscape-Model (van den Broek, Young, Tzeng & Linderholm, 1999). In four experiments, subjects read the texts, sentence by sentence, at their own pace. We first collected activation measures of predictive inferences meaning in a lexical decision task. Second, we collected plausibility evaluation of the predictive inferences with a judgment task: Subjects had to judge if the predictive inference expressed in a sentence corresponded to an event that will probably happen next according to the text they read.

In experiment 1 (cf. Table 1), the activation measures were obtained by comparison with a condition which controlled the effect of lexical priming by text words. According to the data, the predictive inferences were activated to a same level when they were only associatively related to the content of a contextual sentence that preceded a predictive sentence (Associative version) or when they were causally related to the content of that contextual sentence (Causal version). On the contrary, the predictive inferences were not activated when the predictive sentences were preceded by a contextual sentence with words unrelated to the predictive inference (Simple version). Thus the predictive inferences appeared equally activated when supported by the same amount of semantic associations. However, the judgment task, performed after the reading of all the texts, indicated that the plausibility of the predictive inferences was evaluated as higher for the Causal than for the Associative versions. Thus, the amount of causal relations between the text events and the predictive inference might determined the integration of these inferences in readers episodic representation of the text.

In Experiment 2 (cf. Table 2), subjects performed the judgment task after the reading of each text plus three neutral sentences which followed the predictive sentence. The latencies of correct responses were longer for the Associative versions than for the Causal versions. That result again suggests that the predictive inferences were not strongly integrated to readers causal representation of the text events, contrarily to what had already happened when the Causal versions were read.

Experiment 3 (cf. Table 3) and Experiment 4 (cf. Table 4) respectively replicated the results of Experiment 1 and 2 and also showed that the predictive inferences were activated when only the contextual sentences of the former Associative and Causal versions were read (C-Associative and C-Causal versions). However, the judgment data indicate that the predictive inferences were not considered as plausible continuation of the C-Associative versions. Thus, when the contextual

sentence were processed alone and were associatively but not causally related to the predictive inferences, these inferences were activated but presumably not integrated in readers' situation model (C-Associative versions). On the contrary, the predictive inferences were not only activated but presumably integrated in readers' situation model when they were causally related to text information (Causal Contextual version). Thus, applied to predictive inference generation, the distinction between activation and integration processes gives a coherent explanation of the data. Although clearly distinct from a lexical priming effect, a predictive inference would be transitorily activated via semantic associations (Keefe & Mc Daniels 1993) and integrated to readers representation of the text content (Klin, Guzman & Levine 1999). Our result allowed to precise that the activation level of predictive inferences mainly depends on the strength of semantic associations with the inference that are provided by text words. Moreover, the causal relation that can be establish between the associated text information and the inference are not crucial for this activation processes. When these relation are missing, the predictive inference is probably activated as a general expectation. However, our judgment data suggest that the predictive inferences can be integrated at various degree in readers episodic representation of the text content. They also suggest that the integration processes are governed by the causal relation that related the text events to the inference in the situation described by the text.

Table 1. Mean latencies (ms) of positive responses in the lexical decision task (L. D.) and mean percentages of positives responses in the judgment task collected for each text version.

	Neutral	Control	Simple	Associative	Causal
L.D.	688	677	675	645	654
Judgment	14	26	76	75	84

Table 2. Mean percentages and mean latencies (ms) for positive responses in the judgment task collected for each text version.

	Control	Simple	Associative	Causal	Explicit
Percentage	19	68	76	77	90
Latency	-	-	2512	2205	1945

Table 3. Mean latencies (ms) of positive responses in the lexical decision task (L. D.) and mean percentages of positives responses in the judgment task collected for each text version.

	C-Simple	C-Associative	C-Causal	Simple	Associative	Causal
L.D.	658	636	627	649	624	639
Judgment	46	45	77	80	81	85

Table 4. Mean percentages and mean latencies (ms) for positive responses in the judgment task collected for each text version.

	C-Simple	C-Associative	C-Causal	Simple	Associative	Causal
Percentage	40	36	85	89	80	90
Latency	-	-	1479	1489	1411	1337

Inference Generation During Comprehending Written Directions : Example of a Posology

C. Kohler, C. Kékenbosch & J. C. Verstiggel

University of Paris VIII (France)

ckohler@voila.fr

It becomes apparent that the comprehension of written directions must be conceived of as involving a different processing according to the reading goal. For example, cognitive activities brought into play during comprehension involve a preferential processing of the semantic structure whereas those developed during reasoning require the construction of a specific situation model (Kintsch, 1994; Perrig & Kintsch, 1985; Schmalhofer & Glavanov 1985). In the latter case, information provided in the text is matched with background knowledge and integrated with it (van Dijk & Kintsch, 1983). One of the special features of written directions is that the instructions are conveyed in a general and condensed way, even for a particular task. This is based on the fact that written directions generally define macro-actions or general categories supposed to cover all the cases that could be encountered. This is especially apparent with pharmaceutical instructions. Indeed, they often express rules that can be applied to categories of sick persons. Theses rules correspond to a dosage of the drug and to frequencies of application. Generally, they are expressed in a form of intervals. In order to elaborate a particular situation model, one needs to translate intervals given in the text into the form of individual-dose amounts of drug. This translation, which requires the generation of many inferences, leads to the definition of the result to reach. According to Richard (1990), this stage of information processing corresponds to the elaboration of a declarative representation. This is the basis on which one constructs a procedural representation of actions to bring into play. However, the conversion of the representation of the result to attain into "effective procedures" is not direct (Georges, 1988): It is frequently necessary to produce inferences in order to understand and to act on the situation, to define plans, and to elaborate action sequences. Our aim was double : a) To analyse the process underlying the construction of a declarative representation during comprehension instructions in order to use a drug ; b) to study the nature of the required representation for the planification of actions which are necessary to accomplish the task. In order to reach these two aims, we manipulated the reading goal and asked subjects to think-aloud.

Method

All subjects (N = 26) received the same experimental text: The dosage and the directions for the use of a drug. The only difference between the groups laid in the reading goal given in the experimental instructions. In a first group (n = 8), subjects were asked to read and to understand in order to relate what they had read. In a second group (n = 8), subjects were asked to read and to understand in order to prescribe the drug to a depressively ill patient. Furthermore, we asked these two groups to think aloud, that is, to verbalise anything that came to mind when reading a sentence. In a third group (n = 10), we asked a) to read and to understand, b) to carry out the instructions and, simultaneously, c) to verbalise every action brought into play. The verbalizations achieved by the subjects were recorded on a tape recorder. This allowed us to collect verbal protocols from which we elaborated a classification of inferences EFFECTIVELY generated according to the reading goal and supposed to play a part in the construction of each representation.

Results

When it is question to read in order to relate back or to read in order to prescribe, subjects expressed verbalizations which testify the generation of bridging-inferences. We have also found verbalizations which attested that all subjects of these two groups generated micropropositional inferences. These sorts of inferences, elaborated from world knowledge, enabled the specification of textual information and are inherent in the nature of the text itself. Moreover, subjects explained constraints given in the instructions through the generation of causal inferences. Only subjects who read in order to prescribe generated superordinate goal inferences. This kind of inferences is closely linked to the goal that motivates reading: To prescribe the drug (in order to treat). All subjects who read in order to prescribe or to carry out instructions generated particularization inferences. For example, from an information like "4 or 6 tabloids 2 or 3 times daily", subjects inferred that the appropriate dosage for the patient was "2 tabloids in the morning, 2 tabloids at noon, 2 tabloids in the evening". Constructing these inferences implies that relevant information would insert into pre-existent slots of pragmatic knowledge possessed by the reader. Subjects of these two groups generated pragmatic inferences. These are also generated from schemata, but, unlike particularization inferences, they are not directly linked to textual information. Therefore, they seem to depend on an implicit reading goal: To give the application method of the drug. Like particularization inferences, new information indicates the intended result, but it provides no indication about actions to be brought into play in order to satisfy this result. Only subjects who read in order to carry out instructions verbalized inferences related to the planification of actions, the role of which is to specify the overall actions to perform in order to attain the goal. The structure and the content of verbalizations attested to the programming of actions required for the fulfilment of pre-requisites (i.e., actions that must be carried out first in order that the actual procedure could take place) and the construction of sub-goals in working memory. Also, as soon as a sub-goal was defined, subjects planed the sequence of actions to perform in order to satisfy it. We noted that particularization inferences and inferences about the planification of actions come from the same linguistic data; this allows us to make the following two observations: (a) subjects who read in order to carry out instructions elaborated two distinct inferences from the same information; (b) depending on whether the goal was to prescribe or to carry out instructions, the inference constructed from a same textual data depends upon a different point of view.

Discussion

The originality of this experiment is that it permits to test experienced inferences, namely, inferences that are not created by the experimenter, but inferences really identified in subjects verbalizations. It allows us to specify differences in inferential production and in the mode of construction of the representation, which depends either especially on semantic structure or on a situation model (Kintsch, 1994). More precisely, comprehension / memorisation activities involve a preferential processing of the semantic structure of the text. Here, inferences required for textual

information comprehension and for the establishment of coherence are primarily generated. When the goal is to prescribe a drug, subjects construct a situation model in the form of a declarative representation. Inferences enable to derive the result to attain from information given in the text. When the goal is to perform actions, subjects construct a procedural representation on the basis of a declarative one. Also, they elaborate inferences about goals to reach and action sequences to bring into play in order to satisfy these goals. Comprehension of a procedural text can be conceived as a variable depth process. In this sense, the procedural representation of actions to perform is based on a declarative representation of the result to attain, which is itself based on the semantic structure of the text.

How Does Negation Affect Inferences During Reading?

R. B. Lea* & E. J. Mulligan**

*Macalester College (U.S.A.)

**University of Colorado (U.S.A.)

lea@MACALESTER.EDU

Recent research has shown that negation can automatically suppress the activation of concepts in a reader's mental representation of a text (e.g., MacDonald & Just, 1989; Gernsbacher & Jescheniak, 1995). But how does negation affect inference making? In three experiments, we investigated the role that negation plays in the making of deductive inferences during reading. According to at least one account of logical competence, people are able to draw certain propositional-logic inferences easily and accurately (Braine et al. 1995). This claim has held for performance on elementary logic problems, and recently it has been extended to the domain of discourse comprehension (Lea et al. 1990; Lea, 1995). The present work examined two logical inferences: One that yields a positive conclusion (or-elimination: a or b; not a; therefore b) and one that leads to a negated conclusion (not-both elimination: not both a and b; a; therefore not b). According to data collected by Braine et al. (1984) these two inference forms were made equally easily when participants were working with abstract content, but the question arises as to whether the effect that negation has on propositions goes beyond explicit negation. That is, does negation inhibit or suppress the production of negative information?

In Experiment 1 we presented texts in which participants could make the or-elimination inference or the not-both inference. In addition, four control conditions were included in which no inference could be made. Word pronunciation (naming) response times were collected for items related to the inference concept. For example, in a passage about a student who is deciding whether to take History or English, and then later decides not to take English, the inference concept HISTORY was presented as the naming probe. In the not-both version of this text the decision about which course to take is presented in the form of someone who cannot take both History and English, and who later decides to take English. In this case the inference is a negated proposition - that she did not take History. Verification times to comprehension questions about the protagonist's decision also were collected to provide converging evidence for inference making. Both naming and verification results indicated that readers were making the or-elimination inference, though these measures were mixed for the not-both inference. That is, the naming times showed no priming for the inference concept, while the verification times indicated that the inference had been made during reading of the text. This pattern is consistent with the possibility that readers were making the not-both inference while reading, and that the negated inference concept was suppressed, as MacDonald and Just (1989) had found in a naming task of explicitly negated concepts.

An alternative explanation for the results of Experiment 1 is that a change in discourse focus - not inference making - produced the naming time results. In other words, when the naming probe appeared, the discourse focus was shifting toward the probe concept in the or-elimination case but not in the not-both case; Perhaps the faster naming times in the "or" stories simply reflects the discourse focus at the time of test. So in Experiment 2 we presented the premises in reverse order so that no change in focus occurred at the time when the inferences could be made. For example, the passage about a student deciding which course to take was changed so that she began knowing that she had to take English, and then later discovered that she couldn't take both English and History. The inference is that she wouldn't take History. The pattern of results was the same as Experiment 1, thus eliminating discourse focus as the critical factor. As in Experiment 1, the naming and verification latencies were mixed with regard to the not-both inference. Specifically, the naming times for not-both inference concepts were not different from controls (suggesting that no inference had been made), but the pattern of verification latencies indicated that the participants had indeed drawn the inference. Again, this paradoxical result is consistent with the hypothesis that the inference (e.g., not History) was made less accessible because it was negated.

In Experiment 3 we examined reading times for sentences that contained information that was inconsistent with the inference concept. Reading times were a logical choice because previous research has shown that negation does not seem to influence reading time (e.g., MacDonald & Just, 1989). In Experiment 3, we composed passages in which certain target sentences would make sense only if the inference in question was NOT made, but would appear contradictory if the inference was made. This technique has been exploited to investigate a variety of text processing phenomena (e.g., Albrecht & Myers, 1995; Lea, Mason, Albrecht, Birch, & Myers, 1998). For example, in a passage about two brothers making Chili, the recipe states that "they could not use both fresh jalapenos and ground pepper in the chili." When they decide to use fresh jalapenos, the reader can infer that they will not use ground pepper. The target sentence reads "I'll get the ground pepper said Joseph", which is inconsistent with the inference. The only difference between the inference version and the no-inference control version of the passage is the omission of the word "not" in the second sentence. The local coherence of the target sentence is identical (and reasonably good) in both the inference and control versions. The target sentence is only inconsistent when the not-both inference is made. Thus, if readers are in fact making the not-both inference, their reading times should be slower on the target sentence compared to the no-inference control. The reading time data from Experiment 3 showed a very large slowdown on inference passages, which clearly indicates that participants were making the not-both inference. Reading times on the pre-target sentence ruled out a spill-over explanation of the results.

Together, these three experiments present a complex and potentially interesting picture of the role that negation might play in inference making. First, negation does not seem to inhibit the inference making process- at least not in a way relevant to the current inferences; Experiment 3 established that readers were making the negated not-both inference. Second, it appears that negation does affect the inference after it is made. Specifically, the present data indicate that inference concepts that are negated are less accessible than comparable inferences that are not negated. This result is parallel to MacDonald and Just's (1989) finding that explicitly negated concepts (as opposed to inferred negated concepts) were less accessible from the readers mental representation of the text. MacDonald and Just explained their effects in terms of discourse focus. However, there appears to be more going on in the present passages than just focus shifts. By separating focus shift from presentation of the second premise, Experiment 2 established that the priming of the or-elimination inference targets was not due to a shift in discourse focus. In addition, Experiment 3's demonstration that readers were in fact making the not-both inference shows that the inference mechanism takes precedence (at least chronologically) over whatever mechanism drives focus shift. In other words, the inference is made first, and then the focus shift, or suppression mechanism, is activated. Thus, the inference is made even though it yields information that is not particularly relevant to the discourse focus. These results fit well with discourse comprehension theories such as Gernsbacher's Structure Building Framework (e.g., Gernsbacher, 1990), and Kintsch's Construction-Integration model (Kintsch, 1988). The present work also further establishes logical deductions as part of the common repertory of inferences that people make in everyday cognitive processing.

Predictive Inference and Text Characteristics : Can Different Types of Causality Lead to Different Degrees of Activation During Reading ?

A. Teisserenc & P. Maury

University of Paul Valéry (France)

a.teisserenc@caramail.com

It is traditionally accepted that text comprehension is a constructive process which results in the elaboration of a global representation of what the text is about. The construction of this representation requires the activation of both the concepts explicitly mentioned in the clauses and implicit information derived from general world knowledge. This implicit integration during reading is based on inferential processes. In our study we focused on inference activation and especially on predictive inferences. They refer to the on-line activation of consequences of events or actions described in the text. Predictive inferences are usually not thought to be activated on-line because they are not required to make the text coherent and to understand it. However, recent studies showed that they can be activated on-line in certain circumstances (Keefe & McDaniel, 1993; Whitney et al., 1992). Several authors (Graesser et al., 1994; McKoon & Ratcliff, 1992) proposed that their activation can be viewed as a continuum: Predictive inferences are activated at different degrees depending on whether the context is strong or weak. In this view, we wanted to show that text characteristics can lead to different degrees of activation. We worked with two types of context based on two types of causality: Intentional causality vs physical causality. Type of causality is seldom controlled in studies. We thought that this lack of control can explain divergent results. First of all, we hypothesized that a stronger context provided by the intentional causality can lead to reinforced the probability of occurrence of consequent events: In this way, the inference degree of activation would be greater than in a weaker context provided by simple physical causality. In fact, the consequent event is caused and expected by an acting agent in the first case, and only caused by a physical agent in the second case. Our aim was also to confront different methodologies involved in predictive inferences study. We first tested inference activation in two ways: On-line and off-line. We also used two different on-line tasks: A word completion task and a lexical decision task.

Method

Two types of text are used for the whole experiments. They describe scientific mechanisms about physic phenomenon (physical causality) and objects fabrication (intentional causality). In this way we have a human agent specified in the texts concerning the making of objects (paper, plastic, glass) with the intentional idea compared with texts without any intention (i.e. any human agent) about physical, natural mechanisms (lightning, erosion, volcano). Seven experimental texts are constructed in the same way: 3 introductory sentences and 2 sentences with the same agent as the grammatical subject (human agent: chemist, artisan, workman vs natural agent: water, lightning, rocks, waves) - the last sentence leading to a predictive inference. The 9 inferences are first tested off-line in a protocol task in which 70 subjects are told to read the texts and to answer

the question "What happened next ?" which was placed after each sentence leading to a predictive inference. In order to test inference activation during reading two different tasks were used. In the first experiment, we employed a completion task which involves providing the first two letters of the inference target word. Readers were told to complete those items with the first verb that comes to mind and begins with those letters. If the predictive inference is activated on-line, the complete verb would be the one that corresponds to the inference.

The second experiment used the same exact method as the first, but a lexical decision task is used. The same verbs are presented to readers who were told to decide if they are French verbs or not. Three versions of each passage were created in order to evaluate inference degree of activation. Explicit and implicit versions were identical except for the last sentence verb: The inference verbs were provided in the explicit passage. A third control version was created with the same words and context but without leading to the inference. This control version was used for the lexical decision task. For the first task, baseline rate for completing inference verbs were integrated as control condition: 20 other subjects were asked to complete the items without reading the texts.

Results and Discussion

Results from off-line and on-line experiments are presented and compared in order to debate on our interest: Predictive inferences. Predictive inferences were activated in the off-line protocol. A qualitative effect of type of context was obtained: Subjects answers reflected the two different causalities (intentional and physical) through goal markers as connectives in one hand and probability markers on the other hand. The type of context had however no effect on predictive inferences on-line activation: The expected facilitation effect of intentional contexts was not obtained both in the completion task (F (1, 60) = 1.43, p < .23) and in the lexical decision task (F (1, 56) < 1). Completion scores were too low to conclude that predictive inferences were activated on-line, and the difference between implicit and control versions only tend to be significant (F (1, 60) = 2.8, p < .09). In the same way, the difference between lexical decision times in the implicit and control versions was not significant (F(1, 56) < 1). Predictive inferences did not seem to be activated on-line. The comparison between results from off-line (protocol) and on-line (completion and lexical decision tasks) methods of study led us to enhance several problems: First, compared with completion task, our lexical decision results seemed to support the existence of post access effects involved by this task and already criticized by many authors (Potts et al., 1988; Whitney et al., 1992). We saw that predictive inferences can be made off-line when they are explicitly incited: In general reader's answers at the question "What happen next ?" provided different words to express a same consequence. We selected the more frequently used word as the target-verb for the two on-line tasks. But we wonder whether the inference can be reduced to a single word. This question leads to a larger debate concerning the nature of inferences and text representation. First of all, with reference to McKoon & Ratcliff (1986) position, we wonder if predictive inferences can be activated on-line but in a minimal way without corresponding to the exact presented word-item. In the same way, Kintsch (1988, 1998) suggests that reading leads to the build of a propositional representation of text: So we have to wonder if the proposition corresponding to the inference which could be activated on-line is really reflected by a single verb. Fincher-Kiefer (1993) also suggested that predictive inferences are integrated during reading at a more abstract level of representation: The situation model. Because of the variety of words used by subjects to express the inferences, our results from verbal protocol support the fact that they could be made in such a more abstract format. Kintsch (1998) also proposed that the situation model might be an image in certain circumstancies. In this case, the presented verb could not exactly and completely represented the created image during reading. By consequent, the inference could be made on-line without facilitating the treatment of the verb presented for the tasks.

Genre of the Text and the Activation of Elaborative Inferences: A Cross-Cultural Study Based on Thinking-Aloud Tasks

J. A. Leon *, I. Escudero* & P. van den Broek **

*Universidad Autónoma de Madrid (Spain)

**University of Minnesota (U.S.A.)

joseantonio.leon@uam.es

During the last two decades, there has been significant research on inference generation during reading. Whereas many investigators have used speeded techniques, in which one element is probed for activation during reading (e.g. Graesser, Singer & Trabasso, 1994; see Haberland, 1994 or Singer, 1994 for a review), others have chosen verbal protocol methods such as question answering (e.g. Graesser, 1981; Graesser & Clark, 1985; Long, Oppy & Seely, 1997) and think-aloud tasks (e.g. Trabasso & Suh, 1993; Trabasso & Magliano, 1996; Zwaan & Brown, 1996; see Pressley & Afflerbach, 1995, for a review). In think-aloud research, readers have been found to engage in a variety of cognitive processes during reading, including strategic inference-making, text integration, and text interpretation (Pressley & Afflerbach, 1995). In particular, think-aloud methods have revealed inferences and mental processes that occur during comprehension (Olson, Duffy, & Mack, 1980; Fletcher, 1986; Suh & Trabasso, 1993; Trabasso & Magliano, 1995, 1996; Narvaez, van den Broek & Barrón, 1999). In addition, think-aloud tasks show the information available in working memory as well as the information that can be used strategically and is codable in language (Ericsson & Simon, 1994).

The purpose of this study is twofold. On the one hand, we analyze the effects of text type (expository in contrast to narratives and newspaper articles) on the expectations and inferences activated by university students during a think-aloud task. Traditionally, research has focused almost exclusively on two types of text - narrative and expository. The results of this research have been mixed. Some authors suggest that there are similarities in the processing of narratives and expository texts (e.g. Goldman & Varma, 1996) and literary texts (van den Broek, Rohleder, & Narvaez, 1996), whereas others propose that readers spontaneously set in motion different patterns of activation or inferences according to the type of text (Brewer, 1980; Harris et al., 1998; Mc Daniel, Einstein, Cunay & Cobb, 1986; Mc Daniel, Owen & Cote, 1990). For example, some researchers have suggested that narratives are more likely to generate predictive inferences and that expository texts lead to more text integration (Olson, Mack & Duffy, 1981). A second objective is to determine if there is a universal cross-cultural inference-making process or if linguistic differences produce varying inference patterns. There exists very little research on this issue. A few researchers have attempted to analyze cultural differences in the context of telling stories (Chafe, 1980; Tannen, 1980, 1983, 1984, 1988). Other authors investigated memory for narratives by North-Americans and Liberian (children and adults, literate and non-literate) observing similar behaviors (e.g. Mandler, Scribner, Cole & De Forest, 1980).

Most constructionist frameworks adopt the idea that a finite set of inference categories can be produced during comprehension (e.g. Graesser et al., 1994; van den Broek, Fletcher, & Risden, 1993). Trabasso and Magliano (1996) proposed three types of inferences that could be activated during the comprehension of a sentence: (1) Explanations, (2) predictions, and (3) associations. In addition, they attempted to relate the various types of inferences to two other important aspects of comprehension: Sources of information and memory operations. With respect to the first, an inference can be based on the text or on relevant world knowledge. With respect to the second, inferences are related to three memory operations that make information available for inferences. Thus, information can be a) maintained in working memory across story sentences (e.g. for

explanation inferences), b) retrieved from a long-term memory representation of the text (e.g. for predictive inferences), or c) activated from general knowledge (e.g. for association inferences). We hypothesize that each of these memory operations is not necessarily linked to one type of inference but rather than memory operations and inferences are dynamically interconnected.

Method

A total of 40 subjects (20 North American and 20 Spanish university students) took part in this research. We used 20 short texts of 4 different types: 3 fairytales, 6 complex stories, 4 newspaper articles, and 7 expository texts. The texts had been taken from newspapers, novels, encyclopedias, handbooks, and stories used in previous experiments by other researchers. The texts were presented to North-Americans and Spanish students in English and Spanish respectively. Subjects performed a think-aloud task as they read the texts. To categorized the subjects´ responses we used the categories proposed by Zwaan & Brown (1996) and Narvaez, van den Broek and Barrón (1999). All the texts were calibrated in length and causal structure according to the model proposed by Trabasso and Magliano (1996). No explanation was given regarding the kind of texts that the subjects were to read. Texts were presented in random order.

Results

Genre was found to be an important factor in the determining the kinds of inference produced ($p < .001$). The main findings are the following. First, it influenced the generation of explanatory inferences: Expository texts required a greater number of explanations than did newspaper articles ($p < .05$), whereas complex stories required fewer explanations than did fairytales ($p < .05$). No other differences in the number of explanatory inferences generated in the various text types were significant. Second, text generation influenced the generation of predictions: Complex narrative texts gave rise to a significantly greater number of predictions than did newspaper articles and expository texts ($p < .001$), with expository texts eliciting fewer than newspaper articles ($p < .05$). No other differences in predictive inferences were significant. Third, newspaper articles and expository texts produced a greater number of comprehension problems that did narrative texts ($p < .05$). No significant differences amount text type were observed for other inferential activities. Also, no differences were observed between North-American and Spanish students.

Conclusion

In summary, expository texts evoked more explanatory inferences and comprehension problems than other text types, whereas narratives gave rise to a larger number of predictive inferences. Newspaper articles elicited a pattern of inferences that held the middle between those for expository texts and the narratives. Thus, different types of text evoke different working memory operations (Trabasso & Magliano, 1996). Expository texts require more activation of relevant knowledge, more maintenance of information in working memory, and perhaps when this is incomplete, more comprehension problems could be produced. These patterns can be explained in terms of differences in reader's prior knowledge, in structure of the various text types, and in the nature of the contents of the text itself. These results also support the idea that the subjects implicitly realize the kind of text that they are reading. In fact, no explanation was given regarding the kinds of texts that the subjects were to read. This represented a significant departure from previous experiments (e.g. Narvaez, van den Broek & Barron, 1999). The fact that these pattern were observed in both North-Americans and Spanish students indicates that the inference-making process is not restricted by linguistic and/or cultural differences.

Individual Differences in Instrument Inference: Lexical Decision versus Picture Naming

S. Kim* & H. J. Yoo**

*Kwangwoon University (Korea)

**Sung-Kyun-Kwan University (Korea)

sungkim@daisy.kwangwoon.ac.kr

Whether the instrument inference was generated during comprehension has been the controversial topic among discourse researchers. Although Kim, Lee, Lee, and Lee (1998) provided a convergent evidence for the on-line generation of instrument inference, recent studies showed that individual differences in working memory span resulted in differential inference processing (Kim, Lee, Kim, & Yoo, 1999). This study was conducted to investigate the individual differences in instrument inference using lexical decision task (Experiment 1) and picture naming task (Experiment 2).

Method

Participants. 24 undergraduate students (10 low span readers and 14 high span readers) in Experiment 1 and 30 undergraduate students (14 low span readers and 16 high span readers) in Experiment 2 were selected on the basis of their scores on the reading span test. Low span readers were defined as those whose reading span score was 2.0 or less and high span readers 4.0 or more.

Design. 2 (high span, low span) x 4 (explicit, implicit, alternative, control) mixed design was used.

Materials. Reading span test: The materials for the reading span test were based on Daneman and Carpenter (1980). The total of 175 unrelated sentences were divided into 7 sets, with five trials at each span size. Each sentence ranged in length from 7 to 10 words. The reading span score was defined as the highest set size for which the participant correctly recalled all of the words from three out of five sets.

Experimental Materials. A set of 32 experimental sentences was constructed with four versions each. In order to select appropriate instrument in each sentence context, forty-seven students not involved with the experiments were asked to generate the most plausible instrument given particular sentential context. From a list of 61 sentences, 32 instrument-sentential context pairs were selected. The average generation rate of the selected instruments with particular sentential context was 92%. In explicit condition, the instrument was explicitly mentioned in the sentence whereas the instrument was not mentioned in implicit condition. In alternative condition, the improbable instrument was mentioned instead of probable instrument. The control sentence was constructed in order to equate the effect of word-based priming based on semantic association among target words and predicates or other words.

Procedure. Participants read sentences that were presented one word at a time on the center of a computer monitor with RSVP rate of 350 ms plus 15ms per character. In Experiment 1, the participant's task was to decide as rapidly and accurately as possible whether the target word was

a Korean word or not. In Experiment 2, participant's task was to say the name of the picture out loud as soon as the picture appeared.

Results and Discussion

In Experiment 1, a 4 x 2 ANOVA revealed a significant main effect for experimental condition, $F_1(3,126) = 23.65$, $p < .001$, $F_2(3,93) = 10.66$, $p < .001$. Although the main effect for reading span was not significant, the interaction effect for reading span by experimental conditions was significant, $F_1(3,126) = 3.80$, $p < .05$, $F_2(3,93) = 2.59$, $p = .058$. For high span readers, the lexical decision times for implicit condition were significantly faster than alternative and control condition whereas there was no significant difference among the three conditions for low span readers. In Experiment 2, it was found that significant main effects for both experimental condition, $F_1(3,90) = 44.52$, $p < .001$, $F_2(3,93) = 34.33$, $p < .001$, and reading span, $F_1(1,30) = 7.19$, $p < .05$, $F_2(1,31) = 47.40$, $p < .001$. The naming times for explicit and implicit conditions were faster than alternative and control conditions for both high and low reading span readers and low span readers were significantly faster than high span readers. There was no significant interaction effect between reading span and experimental conditions.

The results of Experiment 1 indicated that low span readers did not generate instrument inference on-line whereas high span readers did. However the results of Experiment 2 showed that both high span readers and low span readers generated instrument inference on-line. These contradictory findings can be resolved by differential sensitivity between lexical decision task and picture naming task. It is suggested that lexical decision task is an appropriate task for detecting linguistic inferential processing while picture naming task might be appropriate for reflecting perceptual aspect of inference. Since reading span reflects some kind of verbal ability or processing of phonological loop in working memory, low reading span readers might have some difficulty to generate verbal instrument inference. On the other hand, since picture naming task taps perceptual aspect of instrument inference, which is not directly related to the reading span, low span readers generate perceptual inference on the instrument.

Table 1. Mean Response Times in Lexical Decision Task and Picture Naming Task (ms)

Span	Explicit	Implicit	Alternative	Control
Experiment 1 (Lexical Decision Task)				
High	716 (40.5)	811 (55.3)	871 (57.3)	922 (59.2)
Low	613 (40.5)	716 (55.3)	710 (57.3)	713 (59.2)
Experiment 2 (Picture Naming Task)				
High	728 (30.9)	894 (37.6)	991 (39.8)	1009(41.2)
Low	673 (30.9)	753 (37.6)	834 (39.8)	881 (41.2)

References

Daneman, M. A., & Carpenter, P. A. (1980). Individual differences in working memory and reading. Journal of Verbal learning and Verbal Behavior, 19, 450-466.

Kim, S., Lee, J. M., Kim, S. Y., & Yoo, H. J. (1999). Individual differences in discourse comprehension: The interaction between reading span and text structure. Society for Text and Discourse: Ninth Annual Meeting of the Society for Text and Discourse, Vancouver, B.C., Canada.

Kim, S., Lee, J. M., Lee, J. H., & Lee, K. (1998). Does instrument inference occur on-line during reading? Society for Text and Discourse: Eighth Annual Meeting of the Society for Text and Discourse, Madison, WI.

Hemisphere Differences in the Processing of High and Low Constraint Predictive Inferences

S. Virtue, T. Linderholm & P. van den Broek

University of Minnesota (U.S.A.)

virt0004@tc.umn.edu

The results of several studies suggest that the right and left hemispheres may process inferences differently. First, studies with right hemisphere damaged (RHD) patients have shown that patients with lesions in the right hemisphere have deficits in inference generation. Some RHD patients have longer response times to inference-related words than to unrelated words when compared to controls (Beeman, 1993). Second, split-visual field studies with normal participants show priming of specific types of information in the right and left hemispheres. The right hemisphere shows priming effects for more distant semantically related words (Beeman, et al., 1994). In contrast, the left hemisphere shows priming effects for more closely related associates. These two studies suggest that the right hemisphere may have an important role in making certain type of inferences during text comprehension. One such type is that of a predictive inference, in which readers generate expectations about what will happen next in the text. The reason that one might expect this type of inference to be processed differentially by the two hemispheres is that a critical parameter in the generation of predictive inferences is the amount of semantic constraint for it. In particular, the degree of causal constraint determines the likelihood of a predictive inference (van den Broek, 1994). For example, predictive inferences will be made more often if the text includes sufficient antecedents. This is also reflected in theoretical models of text comprehension (e.g., the Causal Inference Maker, van den Broek, 1990). Thus, the level of textual constraint may influence the production of predictive inferences.

The aim of the present study is to examine the interaction of causal constraint and the hemispheric location of processing during predictive inferences. Findings from brain damaged studies, split visual field studies, and theories of causal inference generation lead to interesting ideas about how the brain may process inferences. These findings suggest that there could be differences in the right and left hemispheres in the way they process predictive inferences, especially when the level of constraint is manipulated. If this is true, then we hypothesize that the right visual field - left hemisphere (rvf-LH) should be faster in responding to highly constrained predictive inferences. In contrast, the left visual field - right hemisphere (lvf-RH) should be faster in responding to low constraint predictive inferences.

Method

Forty-eight inference items were constructed that consisted of a high constraint and low constraint version. The items consisted of 4 sentences that were constructed so that the predictive inference might be generated after reading the last sentence in the item. For both the high and low constraint items, the target words presented to the participants in the lexical decision task were both words associated with the inference. Forty-eight neutral items, that did not promote these inferences, were constructed as a baseline condition for each hemisphere. The neutral items were paired with the same target words as those presented in the inference items. In addition, 48 filler items that did not promote any inferences and were paired with non-word targets were constructed.

After several practice trials, participants read each item on the computer one sentence at a time. Then, a fixation plus was presented in the center of the screen (for 750 milliseconds) to focus the participants' attention. Next, a lexical decision task was presented to the participants. The probe word or non-word was presented either to the participants' left or right visual field for 176 milliseconds. All words appeared approximately 3.5 degrees of visual angle to the left or the right of fixation. The center of each probe was 3.4 cm from the fixation center, and the inner edge of each word was at least 1.5 cm from the outer edge of the fixation point. Participants were instructed to press the index finger on the middle button of the button box for a "YES" response, and press their thumb or first button for a "NO" response. Also, participants were asked to recall the main idea of some of the items they read in 1-2 sentence in order to assure comprehension.

Results & Discussion

The main variables were visual field (left or right) and constraint (low or high). The dependent variable was response time to a lexical decision task. Only items that were correctly answered in the lexical decision task were included in the analysis. A main effect was found on the constraint variable. As expected, high constraint stimuli evoked faster responses than low constraint stimuli. Importantly, an interaction effect was found between visual field and constraint. A priori contrasts established the location of this interaction. Similar levels of priming were found for the highly constrained items in both the lvf-RH and the rvf-LH. In contrast, the low constraint items showed a greater level of priming in the lvf-RH than in the rvf-LH.

The results of this study suggest that the role of constraints play an important role when readers make predictive inferences. Also, the results show that the hemispheres work together to process different types of predictive inferences. Although the left hemisphere in general is faster at processing predictive inferences, this hemisphere appears particularly efficient at processing predictive inferences that are highly constrained by the text. In contrast, the right hemisphere is more efficient at producing predictive inferences that are less constrained by the text. These results indicate that the two hemispheres may play rather different, and complementary, roles in the comprehension of texts.

References

Beeman, M. (1993). Semantic processing in the right hemisphere may contribute to drawing inferences during comprehension. Brain and Language, 44, 80-120.

Beeman, M., Friedman, R. B., Grafman, J., Perez, E., Diamond, S., & Lindsay, M. B. (1994). Summation priming and coarse coding in the right hemisphere. Journal of Cognitive Neuroscience, 6, 26-45.

van den Broek, P. (1990). Causal inferences and the comprehension of narrative texts. The Psychology of Learning and Motivation, 25, 175-196.

van den Broek, P. (1994). The causal inference maker: Towards a process model of inference generation in text comprehension. In D. A. Balota, G. B. Flores d'Arcais, and K. Rayner (Eds.) Comprehension Processes in Reading. Hillsdale, NY: Lawrence Erlbaum Associates, 423-446.

Multidimensional Situation Model: Evidence for the Existence of Links Between Specific Situational Dimensions

N. Blanc & I. Tapiero

University of Lyon 2 (France)

Nathalie.Blanc@univ-lyon2.fr

This study deals with the multidimensional content of mental representation readers constructed from text. It is now well established that the model readers built reflects the multiple dimensions the situation implied: It has been demonstrated that events reported in a text are connected in memory through space, time, causation, intentionality and protagonists (Zwaan, Langston & Graesser, 1995; Zwaan, Magliano & Graesser, 1995; Zwaan & Radvansky, 1998; Zwaan, Radvansky, Hilliard & Curiel, 1998). In a recent study (Blanc & Tapiero, 1999), we have demonstrated that readers monitor the situational dimensions as a function of their prior knowledge, the task they have to perform and the comprehension phase of the model (i.e., construction and updating). In addition, we have underlined the complexity of the protagonists dimension. We showed that the characters personality traits and emotional reactions have to be considered as separate situational dimensions since readers differentially monitor these two dimensions as they proceed through the text. Issued from these results, we carried out another experiment not only to confirm what we previously observed concerning the role of readers' prior knowledge, task demands and comprehension phase on the monitoring of the situational dimensions, but also to test whether some links can be established between specific situational dimensions.

Method

The material we used consisted in a text that described a play. We modified the experimental text we previously used (Blanc & Tapiero, 1999) to strengthen the characters personality traits and emotional reactions by providing more detailled information. The text was composed of two types of sentence, those that referred to the situational framework (i.e., the scenery of the stage) and those that dealt with the situational dimensions (i.e., space, intention, personality traits and emotion), assuming a differential monitoring of these two types of information (Zwaan &

Radvansky, 1998). In addition, we divided the text into two parts to study how the situation model evolves all along the text, that is from the construction to the updating phase. Moreover, we constructed two versions for the second part of the text, the first one was consistent with the situation described in the first part of the text whereas the second one was inconsistent. Indeed, emotion and intention inconsistencies were introduced. We postulated that the processing of emotional inconsistencies should disrupt the representation of characters personality traits information being contradicted. We also expected that the processing of intention inconsistencies disrupt the representation of the characters movements since, in our study, the intention information corresponded to the description of a movement required to accomplish a specific goal. The reading of each part of the text was immediately followed by an inferences judgment task. Statements referred either to the situational dimensions or to the framework. Finally, to investigate whether readers' prior knowledge determine the monitoring of situational dimensions, subjects were provided with prior information on the situation via an introductory text. Before reading the experimental text, participants had to read a text that described either the scenery of the stage or the characters personality traits. Then, half of the subjects of these two conditions (i.e., scenery vs protagonists) had to draw a map of the scenery of the stage by precisely locating all the objects mentioned in the experimental text, whereas the other half had to summarize in writing the part of the text they just read. These two tasks were used to investigate whether the task demands could determine the situational dimension foregrounded. Then, all subjects performed an inferences judgment task. They had to judge whether the statement that appeared on the screen was true or false in relation to the part of the text previously read. The accuracy and the speed of inference judgments were recorded. Then, subjects had to read the second part of the text (either consistent or inconsistent) and performed the second testing phase similar (except for the material) to the one used after the reading of the first part of the text.

Results

The results of this study contribute to a deep understanding of which factors determine the monitoring of the dimensions implied by the situation, and enable us to demonstrate that some links can be established between specific dimensions. First, our findings are compatible with those of previous studies, the multiple dimensions taking part of the situation model readers construct (Zwaan et al., 1995, 1998). Specifically, we obtained evidence for the fact that the monitoring of the situational dimensions evolves all along the text. In the first part, readers foregrounded information relative to characters personality traits and spatial location, whereas in the second part they focused on characters emotion and intention. Moreover, we observed that readers differentially process and integrate information relative to the situational dimensions as a function of their prior knowledge and the task they had to perform. Consistent with our previous study (Blanc & Tapiero, 1999), participants who already knew the scenery more accurately represented the spatial dimension whereas those in the protagonists condition were more accurate for the intention dimension. Also, we showed that the summary task promotes the monitoring of the emotion dimension whereas the drawing task facilitates the monitoring of the spatial dimension. More importantly, our findings support our assumption that the personality traits dimension shares links with the emotion dimension. Indeed, we noticed that the introduction of emotional inconsistencies in the second part of the text leads to a decrease in performance for the personality traits dimension. Consequently, readers integrated inconsistent emotional information and then encountered difficulties to retrieve the contradicted personality traits information. Moreover, we underlined that intention information is considered as important since participants detected inconsistencies relative to this dimension. Finally, our study confirm that the situational framework should be distinguished from the situational dimensions since readers differentially process these two types of information. First, it appeared that participants focused on the framework information during the construction phase whereas they shift their attention on the situational dimensions during the updating phase. Second, we demonstrated that, unsurprisingly, participants in the scenery condition constructed a more accurate and a more available representation of the situational framework than those in the protagonists condition. Finally, the situational framework

was more accurately represented and more rapidly accessible when participants had to perform a drawing than a summary task.

References

Blanc, N., & Tapiero, I. (1999). Multidimensional situation models. Paper presented at the 11th Conference of the European society for Cognitive Psychology, Ghent (Belgique), 1-4 Septembre.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. Psychological Science, 6(5), 292-297.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. Psychological Bulletin, 123(2), 162-185.

Zwaan, R. A., Radvansky, G. A., Hilliard, A. E., & Curiel, J. M. (1998). Constructing multidimensional situation models during reading. Scientific Studies of Reading, 2(3), 199-220.

The Effects of Readers' Prior Knowledge on the Monitoring of Spatial and Emotional Dimensions

S. Guéraud & I. Tapiero

University of Lyon 2 (France)

sabine.gueraud@etu.univ-lyon2.fr

The main purpose of our study was to investigate the multidimensional aspects of the situation model constructed from text (Zwaan, Magliano, & Graesser, 1995). Several studies have shown that readers represent more particularly information about spatiality, causality, intentionality, temporality and protagonist (Zwaan, Langston, & Graesser, 1995; Zwaan, Radvansky, Hilliard & Curiel, 1998; Tapiero & Blanc, 1998). In our experience, we were more specifically focused on the spatial and emotional dimensions implied in the situation described. Indeed, several researches have demonstrated that spatial dimension is difficult to monitor but, at the same time, appears to be necessary to represent the situational framework. Moreover, although Zwaan et al., (1995) did not take into account the monitoring of emotional dimension in the event-indexing model, Gernsbacher, Goldsmith and Robertson (1992) have demonstrated that readers are able to represent this kind of information, even when the characters' emotional states are not explicitly mentioned. As the situation model results from the interaction between textual information and readers' prior knowledge (Tapiero & Blanc, 1998; Zwaan et al., 1998), we also investigated whether the nature of readers' prior knowledge (spatial vs. emotional) determine the monitoring of situational dimensions. First, we supposed that the situational dimension foregrounded in the model readers construct evolves from the construction to the updating phase. Furthermore, we made the assumption that the processing of the multiple dimensions varies as a function of readers' prior knowledge.

Method

To pursue these issues, we constructed a descriptive text, composed of two parts, the construction phase and the updating phase. Each part contained eighteen sentences among which we distinguished emotional and spatial sentences. In the construction phase, the spatial sentences described the location of characters, the furniture and the accessory, and the emotional sentences described protagonist nature traits. In the updating phase, the spatial sentences involved accessory and characters' movement, and the emotional sentences involved the emotional reactions. To test whether readers monitor these two dimensions during the reading, the study of each part of the text was immediately followed by an inference judgment task. Participants had to judge six spatial and six emotional statements according to the part of the text they just read. In addition, to study the effects of prior knowledge on the processing of emotional and spatial dimensions, an introductory text, composed of nine sentences, was presented before the reading of the experimental text. One dealt with the spatial layout, presenting only the location of furniture (spatial knowledge group) and the other one with the characters' emotional reactions (emotional knowledge group). According to the fact that spatial information is necessary to represent the situational framework (Zwaan & Radvansky, 1998), we predicted that the spatial dimension should be foregrounded during the construction phase whereas readers should focus on the emotional dimension during the updating phase. Moreover, we assumed that the monitoring of the dimensions should depend on readers' prior knowledge (Blanc & Tapiero, 1998). Subjects with spatial prior knowledge should focus on the emotional dimension whereas those who already know the characters' emotional reactions should focus on the spatial dimension. In addition, we expected that participants with spatial knowledge should be more accurate and more rapid to judge spatial inferences whereas those with emotional knowledge should obtain higher performances to emotional inferences for the first inferencejudgment task. On the other hand, the number of correct reponses should be more important for emotional inferences than spatial inferences for the second inferences judgment task, whatever the nature of prior knowledge given to participants. More importantly, we predicted that, during the construction phase, the monitoring of the emotional dimension is easier when readers have spatial prior knowledge than when they do not have prior knowledge on the situational framework implied.

Results

Our results indicated first, that the emotional sentences were longer processed during the updating phase than during the construction phase whereas no difference between these two phases occurred for the spatial dimension. In addition, we observed that the emotional dimension was more accurately represented and more rapidly available than the spatial dimension, whatever the nature of prior knowledge provided (emotional vs. spatial). These findings are consistent with previous researches which have shown that spatial dimension is difficult to monitor compared to others dimensions (Zwaan et al., 1995, 1998) and confirmed that the situation model readers construct from text should be considered as a multidimensional representation. Moreover, two important points go beyond previous studies: First, unexpectively, the spatial dimension is relatively well processed during the updating phase and second, the emotional dimension appears to be a part of the multidimensional representation constructed by readers. Second, we noticed that subjects with emotional knowledge have shorter response times than subjects with spatial knowledge. Thus, emotional prior knowledge appeared to facilitate access to situational information. Finally, our results revealed that the situation model availability varies as a function of readers' prior knowledge and of comprehension phase. Indeed, participants having spatial knowledge showed longer response times during the construction phase than during the updating phase, whereas those having emotional knowledge presented the reverse pattern of response. In others words, spatial prior knowledge seemed to facilite the access to situational information during the updating phase, whereas this facilitation appeared during the construction phase for

subjects with emotional prior knowledge. This finding indicated that the retrieval of situational information depends not only on the nature of prior knowledge but also on the comprehension phase.

Discussion

Our study gives arguments in favour of a multidimensional model of situation constructed from text. Futhermore, it shows the importance of the spatial dimension and especially of the emotional one, in the mental representation elaborated by readers. Therefore, it would be interesting to investigate this emotional dimension in interaction with other ones proposed by Zwaan et al., (1995) or to study the effects of the emotional valence of information (positive versus negative) on the reading processes. At the same time, the processing of the spatial dimension observed during the updating phase leads to rise a question about other factors that could influence this reading process, even if we can already set (following our results) the main influence of prior knowledge on information processing. Finally, as expected, the elaboration of the situational dimensions depends on the nature of prior knowledge (spatial versus emotional) given to readers. Moreover, this effect appears to be modulated by the comprehension phase (construction versus updating).

Cognitive and Emotional Aspects of the Reader's Response to a Story

M. C. Levorato & P. Boscolo

Università di Padova (Italy)

boscolo@uxl.unipd.it

In the 70s and 80s, research on text comprehension has mainly focused on the nature of knowledge representation and its role in reading processes and memory organization. In the last 15 years the research field has widened. On the one hand, increasing attention has been paid to other cognitive aspects of comprehension, such as the reader's inference generation and search for coherence (e.g. Graesser, Singer, & Trabasso, 1994; Graesser, Gernsbacher, & Goldman, 1997; van den Broek, 1997). On the other, several scholars have underlined the limitations of a purely cognitive approach to comprehension and the need for considering the interactions of affective and cognitive components of reading (Miall, 1989; Gernsbacher, Goldsmith, & Robertson, 1992; van Oostendorp & Zwaan, 1994).

Contributions to a wider conceptualisation of text comprehension have come from various approaches, and particularly from research on interest and response to literature. Studies on interest (Hidi, 1990; Hidi & Baird, 1986; Renninger, Hidi & Krapp, 1992; Schraw, Bruning, & Svoboda, 1995) have highlighted the features which make a text interesting (novelty, curiosity, suspense, life themes, etc.) and the effects of interest on learning and memory. Research on response to literature (Beach & Hynds, 1991; Halàsz, 1996; Kreutz & McNealy, 1996; a Rosenblatt, 1983) analyses the reader's interaction with a text, that is his/her involvement, emotional responses, mental representation of characters' intentions and emotions, empathy with characters and the author. By focusing on the reader-text interaction, rather than on the separate aspects of text structure, on the one hand, and reader's processes on the other, this approach truly represents a new frontier in the psychological study of reading. In fact, in order to understand in depth the experience of reading, and particularly of reading literary narrative, the reader's subjective experience and emotions have to be considered and analyzed.

The aim of this study was to analyse the pattern of cognitive and affective-emotional components of adult readers' response to literary narrative reading. A basic assumption of this study was that these components are not separable in reading experience. Emotional states arise from text processing and affect the reader's coding processes and search for information in the text. Thus, they facilitate the comprehension process and make information easier to remember. The hypothesis of a close relationship between a reader's emotional response reading and his/her cognitive operations in reading was synthesised in Brewer's (1996) structural-affective model. Brewer has introduced a major distinction between event structure (the organization and sequence of events in the story) and discourse or text structure (the sequence of events narrated in the text). According to this model different reader's responses (interest, curiosity, surprise, suspense, etc.) may arise from different text structures. Thus, different versions (text structures) of a same narrative text should elicit different patterns of readers' cognitive and affective responses.

This study was aimed at investigating the ways in which the discrepancy between text and event structure affect: (1) Reader's cognitive-affective responses, such as interest, curiosity, surprise, suspense; (2) Readers' evaluation of text comprehensibility and memorability, as well as post-dictability and imagery; (3) Affective responses such as liking, emotionality and involvement. Moreover, a relevant research question concerns the interaction between these three aspects of response to literary reading.

Method

Participants. One hundred thirty-five psychology students participated in the study.

Materials and procedure. Three versions of E. A. Poe's "The oblong coffer", a classic of gothic genre, were prepared. One was the original story, the second was the story with a flash-back which anticipated some crucial events, the third was the original one, with the conclusive sentence, which describes the intense emotional reaction of the narrator, moved up to the beginning. Each version, 12 pages long, was submitted to a group of readers (N = 45), who were asked to read the story and rate it on various measures which are typical of a reader's main response types to a narrative text: (1) Their emotional states (involvement, empathy, etc.) at each critical point ; (2) Their cognitive states (predictability of the ending, imagery, importance of various elements of the story, etc.) ; (3) their affective-cognitive states, that is the states induced by a discrepancy between the available and expected information (interest, curiosity, surprise, etc.). Each of the answers was obtained through a six point scale (1: non at all interesting, comprehensible, etc. 6: very interesting, comprehensible, etc.)

Results

As far as the answers given to the six point scale, a main result obtained through MANOVAs, was a confirmation of Brewer's model, since some significant differences between the responses elicited by the three versions were obtained. However, this result, should be confirmed by an analysis of the reader's response *during* the reading and non only after. A second result regarded the factor analysis of the cognitive, affective and cognitive-affective ratings, from which a complex pattern of results emerged, showing that the three aspects are highly interconnected.

References

Beach, R., & Hynds, S. (1991). Research on response to literature. In R. Barr, M. L. Kamil, P. B. Mosenthal, & P. D. Pearson (Eds.), Handbook of reading research (Vol. 2, pp. 453-489). New York: Longman.

Brewer, W. F. (1996). Good and bad story ending and story completeness. In R. J. Kreutz and M. S. McNealy (Eds.), Empirical approaches to literature and aesthetics (pp. 261-274). Norwood, NJ: Ablex.

Gernsbacher, M. A., Goldsmith, H. H., & Robertson, R. R. (1992). Do readers mentally represent character's emotional states? Cognition and emotion, 6, 89-112.

Graesser, A. C., Gernsbacher, M. A., & Goldman, S. R. (1997). Cognition. In T. A. van Dijk (Ed.), Discourse as structure and process (pp. 292-319). London: Sage.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. Psychological Review, 101, 371-395.

Halàsz, L. (1996). General and personal meaning in literary reading. In R. J. Kreutz and M. S. McNealy (Eds.), Empirical approaches to literature and aesthetics. Norwood, NJ: Ablex.

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. Review of Educational Research, 60, 549-571.

Hidi, S., & Baird, W. (1986). Interestingness: A neglected variable in discourse processing. Cognitive Science, 10, 179-194.

Kreutz, R. A., & McNealy, M. S. (1996)(Eds.). Empirical approaches to literature and aesthetics. Norwood, NJ: Ablex.

Miall, D. S. (1989). Beyond the schema given: Affective comprehension of literary narrative. Cognition and Emotion, 3, 55-78.

Renninger, K. A., Hidi, S., & Krapp, A. (1992). The role of interest in learning and development. Hillsdale, NJ: Erlbaum.

Rosenblatt, L. (1983). Literature as exploration (4th edition). New York: The Modern language Association of America.

Schraw, G., Bruning, R., & Svoboda, C. (1995). Sources of situational interest. Journal of Reading Behavior, 27, 1-17.

van den Broek, P. (1997). Discovering the cement of the universe: The development of event comprehension from childhood to adulthood. In P. van den Broek, P. J. Bauer, & T. Bourg (Eds.), Developmental spans in event comprehension and representation (pp. 321-342). Mahwah, NJ: Erlbaum.

van Oostendorp, H., & Zwaan, R. A. (1994). Naturalistic text comprehension. Norwood, NJ: Ablex.

Updating a Mental Model: On-Line Accessibility of Backgrounded Information

J. D. Murray

Georgia Southern University (U.S.A.)

JMurray@GaSoU.edu

Recently, O'Brien, Rizzella, Albrecht, and Halleran (1998) examined the process by which readers "update" their knowledge of a story's protagonist. In their experiment, participants were timed in reading a sentence (e.g., "Mary ordered a cheeseburger and fries.") that described the protagonist executing a behavior that was either consistent with what was initially stated about the protagonist (e.g., Mary was described as a junk food addict), inconsistent with what was initially stated (Mary was described as a strict vegetarian), or was consistent with a qualification of the inconsistent information (Mary was a strict vegetarian, but occasionally ate greasy food when dining out.). The target sentence was read more quickly in the consistent condition than the inconsistent condition, replicating earlier research (Albrecht & O'Brien, 1993) showing that readers are sensitive to breaks in global coherence. Of primary interest was reading time in the qualification condition, where the target line was read more slowly than in the consistent condition but more quickly than in the inconsistent condition. The researchers interpreted this result as indicating that readers consulted both the inconsistent and the qualification information when interpreting the target sentence.

The focus of the present research was to explore the extent to which backgrounded information that has been updated is accessible later in a passage when protagonist-relevant information is presented. Specifically, is the backgrounded information automatically activated, as would be predicted by a passive, resonance process (Myers & O'Brien, 1998)? Or is it reinstated via a slower, more strategic process? Stated with respect to the example, the present study tested the extent to which "vegetarian" was accessible after reading "Mary ordered a cheeseburger and fries.". This is of primary interest in the case where readers' initial knowledge is subsequently updated (e.g., Mary is described as a strict vegetarian but one who occasionally eats greasy food). Although O'Brien et al.'s results indicate that this information was consulted, it is unclear from their results whether the original information was accessed automatically or strategically.

The present research repeated and extended O'Brien et al.'s experiment. Participants read passages one line at a time and were also timed in naming a word congruent with the inconsistent information (e.g., "vegetarian") immediately after reading the target sentence. With respect to reading times on the target sentence, a replication of O'Brien et al.'s results was predicted: the slowest target sentence reading time should be in the inconsistent condition, followed by the qualification condition, which, in turn, should be slower than the consistent condition. With the naming times, it is predicted that naming times will be faster in the inconsistent condition than in the consistent condition $(I < C)$. This prediction is based on the fact that readers will have had exposure to the inconsistent information only in the inconsistent condition. Of primary interest, however, is the accessibility of the inconsistent information in the qualification condition. It was reasoned that if this information is automatically retrieved when the target sentence is encountered, then naming times in the qualification condition should be faster than in the consistent condition, and similar to the inconsistent condition $[(I = Q) < C]$. On the other hand, if the inconsistent information is not automatically activated but is consulted through a slower, more strategic process, then naming times in the qualification condition would be slower than the inconsistent condition, and similar to the consistent condition $[I < (C = Q)]$. A third possibility is that the naming data will reflect the original reading time data, where the inconsistent, qualification, and consistent conditions are all statistically different from one another $(I < Q < C)$.

Method

Undergraduates enrolled in Introductory Psychology (n = 110) participated in exchange for extra credit. Participants read a total of 24 passages, one line at a time, on a computer display, preceded by three practice passages. The eighteen experimental passages were borrowed from O'Brien et al. (1998). Each passage had three versions, each differing in the relatedness of the target sentence to the backgrounded information pertaining to the protagonist. In the consistent condition, the target sentence was consistent with the backgrounded information. In the inconsistent condition, the target sentence was inconsistent with the backgrounded information. In the qualification condition, the target sentence was inconsistent with the backgrounded information but was consistent with a briefly stated qualification of that information. The qualification version of each of O'Brien et al.'s original items was modified to ensure that participants would perceive the "updated" information as consistent with the protagonist's behavior depicted in the target sentence. This modification was pre-tested in a separate study prior to running the present experiment. As in O'Brien et al., the target sentence was identical across all three versions. Each participant saw six passages in the three experimental conditions. Passages were counterbalanced across condition such that, across all participants, each passage appeared in each condition an approximately equal number of times. Six additional items were filler passages, and basically were similar in form to the items in the consistent condition. These items were randomly interspersed among the experimental passages but were not analyzed.

Each passage was preceded by a "Ready" signal. Participants read each passage at their own pace by pressing a line advance key. Each line of the passage appeared in the center of the computer screen, and was replaced by the next line when the key was pressed. The time between key presses was stored as reading time. Immediately following the target sentence, a warning signal (XXX) appeared in the center of the screen for 500 milliseconds and was followed by a word representing the inconsistent information. Participants pronounced the target word as quickly as possible, into a microphone. This utterance activated a voice-key and resulted in the target word disappearing from the screen. Immediately following, the time interval spanning the onset of the target and the participant's first utterance was briefly displayed on the screen, and was stored as naming time. A second warning signal (???) then appeared for 500 milliseconds and was followed by a comprehension question. Participants responded by pressing a key marked YES or NO. An error message appeared if the question was answered incorrectly. An instruction to read more carefully appeared on the screen after each incorrect response following the fifth error.

Results and Discussion

The primary data were the target sentence reading times and the target word naming times in the three experimental conditions. The data from nine participants were omitted from the analyses. These participants were omitted either because they answered four or more comprehension questions (16.7%) incorrectly or because of problems prematurely tripping the voice key during the experimental session. The analyses, therefore, were based on data from 101 participants. The analysis of the reading times showed, as predicted, that participants read the target sentences more slowly in the inconsistent condition (m = 2487 msec) than in the consistent condition (m = 2221 msec) (p < .001). Also, similar to the findings of O'Brien et al. (1998), reading times in the qualification condition (m = 2282 msec) were slower than those in the consistent (albeit, not significantly so) and faster than those in the inconsistent condition (p < .001). These data are consistent with the notion that, in the qualification condition, participants consulted both the backgrounded, inconsistent information as well as the "updated" qualification information at the point they encountered text material that is relevant to previously read information.

The naming time data fell into line with the automatic retrieval hypothesis. As originally predicted, naming times were faster in the inconsistent condition (m = 449 msec) than in the consistent condition (m = 462 msec), p < .05. However, participants were equally as fast to name the target word in the qualification condition as in the inconsistent condition (m = 449 msec). These data indicate that readers have fast, direct access to backgrounded information of the protagonist, even if that information is inconsistent with an updated mental model of the protagonist. This is consistent with the notion that readers access background information via a passive, fast-acting resonance process (Myers & O'Brien, 1998).

Analysis of Descriptive and Narrative Texts Written by French and Dutch Children

L. Chanquoy* & J. Schilperoord**

*University of Paul Valéry (France)

**Utrecht University (The Netherlands)

chanquoy@bred.univ-montp3.fr

Most writing research shows that writing is a very complex activity, mainly because writers have to manage simultaneously or successively a great number of processes. They have to select and to organize information retrieved from long term memory (planning process), to choose correct syntactic frames and appropriate lexical items (translating process), to reread and to revise the already produced text (revising process) and, finally, to program and execute the necessary motor movements for the realisation of the written text (graphic execution). The study of these different processes has led to the elaboration of models, as Hayes and Flower's (1980) well known model, or more recently, Hayes' (1996) and Kellogg's (1996) models. The research proposed here works from the perspective of these recent models. The objective was to analyze how novice writers (3rd and 5th graders), French and Dutch, managed in real time the processes of producing a written text, and more particularly the recovery of pieces of information, and their syntactic and lexical organization in a text. The results of two experiments have therefore been compared: The first one carried out in French by Chanquoy, Foulin, and Fayol (1990), and the second led with Dutch children by Schilperoord, Chanquoy, and van den Bergh (1997).

Method

These two experiments used, in writing, the method elaborated by Holmes (1984, 1988) for speaking. This method consisted in asking participants to produce an ending to a given text respecting a certain number of constraints. By following the main principles of this method, 3rd and 5th graders were asked to produce text endings following three constraints: The predictability (predictable vs. unpredictable text content), the type of text (descriptive vs. narrative) and the syntactic structure (one sentence with two [or three in Dutch] clauses vs. two [or three] sentences). Thus, these two experiments aimed to study the impact of these different constraints on interclause (or inter-sentence) and pre-writing pause duration, and on writing speed. Each child, French or Dutch, wrote 8 endings, which had to respect the three constraints described above (predictability, type of text, and syntactic complexity). During all the time of their writing,

children were videorecorded (the sheet of paper, the hand and the pencil) with a camera allowing very precise chronometric measures. At the end of the experiment, three measures were collected: The duration of pre-writing pause (time between the end of the instructions and the beginning of the graphic activity), the duration of interclause (or inter-sentence) pauses, that corresponded to the time between the end of the writing of the last word of a clause (or a sentence) and the beginning of the writing of the first word of the following clause (or sentence), and the writing speed (measured in number of characters written per second).

Results and Discussion

The results of the French study showed that syntactic, semantic and textual variables played an important role in the temporal management of writing, both at the level of pause duration (i.e., pre-writing and interclausal pauses) and writing speed. Thus, predictable endings led to shorter pre-writing and inter-clause pause durations than unpredictable endings. Narrative endings had shorter pauses than descriptive endings (this result was however only significant for 5th graders). Pauses were shorter and writing speeds were more rapid for endings composed of one sentence (with 2 clauses) than for those of two sentences. Finally, concerning the grade, 5th graders made shorter pauses and wrote more rapidly than 3rd graders.

The results of the Dutch study showed that Dutch children did not manage their writing, according to the imposed constraints, in the same way as French graders. More precisely, both Dutch graders made significantly shorter pauses when they had to write unpredictable endings than when these endings were predictable; This was significant both for pre-writing and interclause pause durations. However, 3rd and 5th graders wrote unpredictable endings significantly slowly than predictable endings. For both grades, pre-writing pause durations did not significantly vary for endings built with one or three sentences. Conversely, interclausal pause durations were longer for endings with three sentences. Whatever the grade, pre-writing pauses and interclause pauses were longer when the endings were narrative. Finally, 3rd graders made shorter pauses than 5th graders. The comparison of the two experiments showed that:

Concerning the grade, while French 3rd graders made longer interclausal pauses than 5th graders, the opposite result was observed for Dutch participants: 5th graders' pre-writing and interclausal pause durations were longer than 3rd graders' durations. For both populations, the writing speed was slower for 3rd graders than for 5th graders. With respect to the constraint of predictability, and as already mentioned, in French, unexpected endings led to pre-writing and interclausal pause durations longer than predictable endings, while the opposite result was found in Dutch. Moreover, while this variable had no effect on the writing speed in French, the writing speed was more rapid for predictable endings in Dutch. For the syntactic constraint, the results were close for pre-writing and interclausal pause durations in French and in Dutch. However, while the writing speed did not significantly vary in Dutch, it was more rapid for endings built with one sentence in French, compared to the two sentence endings (but only for 3rd graders). Concerning the type of text, no effect of the writing speed was noticed both for French and Dutch children. However, while pre-writing pause durations did not significantly differ in French endings, for the Dutch children these pauses were longer in narrative than in descriptive endings. In addition, an opposition between the two populations was observed for interclausal pause durations: these durations were longer for French descriptive endings (in 5th graders) and shorter for Dutch descriptive endings, compared to narrative endings. Thus, while strong semantic constraints led to an increase of pause durations and to a decrease of writing speed in French children's texts, this phenomenon was not observed for the Dutch children. These latter results were very complex to analyse and therefore we investigated them further, notably by analysing verbal protocols collected only from Dutch children. For example, the finer analysis of Dutch children's ending showed that these children used, so as to palliate difficulties of the task, strategies that had not been observed in French participants. Concerning the predictability constraint, the element of surprise in unpredictable endings was simply brought about by a negation or by an inversion of one or several sentences provided in the

beginnings. This phenomenon could then explain why Dutch children wrote unpredictable endings more rapidly.

Sources of Reading Comprehension Difficulties in Children: On-line Processing of Cohesion in Listening Comprehension

M-F. Ehrlich & H. Megherbi

University of Paris 5 (France)

ehrlich@psycho.univ-paris5.fr

Investigate listening comprehension is a way to specify the nature of the deficits in reading comprehension. Two hypotheses can be considered: The linguistic hypothesis views poor readers as suffering problems in language processing whereas the cognitive hypothesis involves more basic deficits (attention, working memory capacity, processing speed…). Previous investigations have shown that poor readers are deficient in processing cohesion devices such as pronouns in written texts. Their difficulties are stronger in situations involving a high cognitive load (Ehrlich & Rémond, 1997). However, the specificity of these difficulties to the written modality and the factors responsible for them remain an open question. The purpose of the two experiments reported in this communication was to investigate the on-line processing of cohesion in a listening situation, in two groups of children contrasted according to their reading comprehension ability. The experiments studied how children integrate different kinds of information for linking sentences by means of a cross-modal naming task used by Marslen-Wilson, Tyler and Koster (1993) in adults. Subjects heard short texts followed by a sentence fragment which had an anaphor in subject position followed by a verb phrase. At the end of the fragment, a visual probe was presented which subjects had to name as quickly as possible. The visual probe was appropriate or inappropriate. The appropriateness effect on the naming latencies of the visual probe reveals the processing of the anaphor.

Experiment 1. On-line processing of two types of subjects anaphors: Repeated noun phrase and pronoun

Subjects. 106 7-year-old children (first grade) were submitted to a reading comprehension test and to a listening comprehension test (French adaptation of Neale, 1988). From the distribution of reading comprehension scores, two groups were formed: Good readers (N = 31) and poor readers (N = 35).

Material and Procedure. 16 short texts were constructed containing a context sentence and a critical sentence. Children heard the text. The last word of the critical sentence was presented visually. Children had to name this visual probe. In addition, 16 filler texts were presented with a different structure. Example of an experimental text: "After a long trip around the world, Elodie had diner with Sebastien in a restaurant (Context sentence). She glady spoke with (Critical sentence) / (Visual probe).

In all texts, the two protagonists (A and B, here Elodie and Sebastien) in the context sentence differed in gender, the verb of the critical sentence was neutral with respect the protagonists. The visual probe was always an object pronoun (lui ou elle). In the critical sentence, the variables were the following: (1) Type of subject anaphor: Repeated noun phrase (Elodie) or pronoun (she); (2) Subject anaphor: Protagonist A or protagonist B; (3) Visual probe: The probe A (him) is

appropriate with the protagonist A in subject position (Elodie), so the probe A is inappropriate with the protagonist B as subject. The probe B (she) is appropriate with the protagonist B in subject position (Sebastien). So the poste B is inappropriate with the protagonist A as subject. The difference between the naming latencies between probe A and probe B for each subject anaphor (appropriateness effect) reveals the processing of this subject anaphor.

Results. If poor readers suffer a specific deficit in anaphor processing (linguistic hypothesis), differences in the appropriateness effect between the two groups are expected. Naming latencies were significantly longer for poor readers. They also were significantly longer for repeated noun phrases than for pronouns. The reading comprehension ability did not interact significantly with any other factors. However planned comparisons showed similarities and differences in good and poor readers. For repeated noun phrases, good and poor readers showed similarly an appropriateness effect only when they refered to the protagonist A. For pronouns, good readers showed an appropriateness effect when pronoun refered to protagonist A, whereas poor readers tended to show this effect (non significantly) when pronoun refered to protagonist B. In conclusion, poor readers seems to have specific difficulties in processing pronouns, but results are not conclusive.

Experiment 2. On-line integration of lexical properties of the pronoun and semantic verb bias

Subjects. Two groups of good readers (N = 29) and poor readers (N = 33) were formed from a pool of 105 8-year-old children (second grade).

Material and Procedure. The cross-modal naming task was used as in Experiment 1. The experimental texts had the same structure as in Experiment 1, but the subject anaphor was always a pronoun. A new variable was manipulated : Three sets of texts varied as a function of the semantic feature of the verb of the critical sentence. Set 1: The verb was neutral as in Experiment 1. Set 2: The verb biased towards the protagonist A. Given that the pronoun refered to protagonist A or protagonist B, the lexical cue on the pronoun and the verb converged towards probe A in half texts. In other texts, there was a conflict between pronoun and verb bias. Set 3: The verb biased towards the protagonist B. In a symetric way to set 2, pronoun and verb bias converged towards probe B or were in conflict.

Results. Naming latencies were similar for good and poor readers. A significant interaction was observed between reading comprehension ability and the factors underlying the appropriateness effect. In sentences with neutral verb, good readers were able to process the lexical cue of the pronoun as adults in Marslen-Wilson and al's study. Poor readers processed the lexical cue when pronouns refers to protagonist B, but not when it refers to protagonist A. In sentences with verb bias, the appropriateness effect was different for good and poor readers. When the lexical cue of the pronoun and verb bias were convergent, the appropriateness effect was smaller in poor readers than in good readers. When the two factors were in conflict, the weight of the lexical cue of the pronoun was dominant relatively of verb bias, especially in good readers.

To summarize, results supported the linguistic hypothesis : 1- Poor readers have difficulties to process the lexical cue of the pronoun in some conditions. 2- Poor readers do not integrate efficiently lexical cue of the pronoun and semantic inference from the verb.

Conclusion

Difficulties observed in the processing of cohesion in written texts are not specific to this modality. Poor readers are also deficient in understanding spoken utterances and results support the linguistic hypothesis.

References

Ehrlich, M.-F. & Rémond, M. (1997). Skilled and less skilled comprehenders: French children's processing of anaphoric devices in written texts. British Journal of Developmental Psychology, 15, 291-309.

Marslen-Wilson, W.D., Tyler, L.K. & Koster, C. (1993). Integrative Processes in Utterance Resolution. Journal of Memory and Language, 32, 647-666.

Transitory Situation Model and Linear Procedural Model: Their Roles in Comprehension of Word Problems Statments at Primary School

D. Coquin-Viennot & S. Moreau

University of Poitiers (France)

Daniele.Coquin@mshs.univ-poitiers.fr

Two categories of studies are examined: Some mention intermediate representations between the instructions and the schema, others are more centered on the activity of the participant rather than on the formal properties of the problems.

Models which include a situation model: The authors (Reusser, 1989; Nathan, Kintsch & Young, 1992; Stern & Lehrndorfer, 1992; Staub & Reusser, 1995) bring into play what they call the situation model, a more qualitative and less mathematical representation of the text of the problem, in order to construct more efficient representations for solving it. The models constructed by these authors show that for them, taking into account the quantitative elements of the instructions (numerical values and quantitative relations) is not the only thing brought into play in order to solve the problem. They also attribute importance to " the intermediate representations between the body of the text and the mathematisation of the problem " (Kintsch, 1998, p. 347); the activation of the schema is not called into question, but it passes through different paths (Kintsch, 1998).

Contextualist or procedural models: In the second category are to be found models known as contextualist (Bastien, 1997), constructed in a school context (Greer, 1997; Reusser and Stebler, 1997), and which for the purposes of this study will be called procedural since they are centered more on the activity of the participant than on the formal properties of the problem. In these models reading activates more or less efficient chain of rules of actions: For example, Bouchafa (1985) suggested a model for the processing of problems composed of three successive modules: a) extraction of the numerical data in the order in which they appear in the text; b) partition of this set of numbers on the basis of various criteria like " belong to the same set ", or " same size "…;

c) selection of arithmetical operators to place between the numbers on the basis of " semantic triggers " or the activation of elementary schemas. We accept Bouchafa's model with the b) partition-criterion: the criterion of " linear partition "; groupings of data are made by order of appearance, the first with the second, and the combination obtained with the third etc. This criterion is natural, insofar as reading is a linear activity with privileged places for the integration of information (Passerault & Chesnet, 1991). In the case of instruction problems, the information is also treated as the reading progresses, including numerical information (Fayol, Abdi & Gombert, 1987; Reusser & Stebler, 1997).

The objective of the current study was to determine whether the processing brought into play during the reading-solving of problems by 11 years old children belongs to the first category of models or to the second. With this in mind, the pupils thus had to solve problems of the distributive type; for example: *For the school party, the school cook puts on the 13 tables [a basket containing] 6 bananas and 8 apples on each. What is the total number of items of fruit that she will use?* They could be solved by two different strategies: Development (13X6 + 13X8) or factorization 13 (8 + 6). The formulation of instructions varied between two crossed variables: presence/absence of a structuring term such as " basket ", and the order of numerical data. The dependent variable was the choice of the solution strategy. The solution models which give importance to the situation model (first category) enable the effect of the presence of a structuring term, and the absence of order effect; whereas the linear procedural model predicts an influence of item order and no effect of the structuring object.

Method

Participants. 182 pupils from 10 different grade 5 classes in 7 schools from the Poitiers region in France. Average age: 11, s.d. = 0.55.

Material. Three distributive problems, each presented in four versions, themselves the result of crossing two variables: Presence/absence of a structuring element (S+/S-) and the placing of the scalar (O2). In the example above, the structuring element is between square brackets and the scalar (13) is at the beginning of the word problem.

Procedure. The distributive problems were combined with four additional distracting problems, and the experiment was carried out in two sessions, with an interval of about 10 days; each pupil solved three distributive problems under the same experimental conditions, one during the first session and two during the second. The order of the problems was counterbalanced.

Results

Performances. 74% to 81% of the problems were solved correctly (92% to 93% if calculation errors were not taken into account).

Strategies. Each factorisation (correct factorisations, e.g. (6+8) x 13; or incomplete factorisations: factorisations initiated (6+8), but not completed or erroneously completed) was coded 1 and the rest zero and then a total was made across all the problems; an indicator of factorisation was thus obtained, varying from 0 to 3 for each participant.

Factorisation was stronger when there was a structuring element in the text (M = 2.13) than when there was not (M = 1.66), ($F[1, 174] = 4.31$; $p < 0.039$), but the place of the scalar had no effect (m=1.70 and 2.08; $F[1, 174] = 3.01$; $p > 0.08$), and there was no interaction ($F[1, 174] <1$).

According to our hypotheses, these results are compatible with a processing model which takes into account the situation model rather than with the linear procedural model. Despite the absence of interaction, the superiority of the factorisations in the presence of a structuring element disappeared when the scalar was placed at the end of the problem (F[1, 174] <1). However, when there was no structuring element, the effect of the placing of the scalar approached significance (F[1, 174] = 3.66; p < 0.057).

Discussion

According to our hypotheses, the results of the canonical analysis, with as the only significant effect the presence of a structuring element, support the idea of a processing model which brings in the situation model. However, the effect of the structuring element disappears when the scalar is placed at the end; for example when the data order was 6 bananas, 8 apples, and 13 tables, the grouping ((6,8), 13) was observed, which can be explained either by the structuring element when it is present, or by linear processing. On the contrary, when the scalar was at the beginning (13 tables, 6 bananas, 8 apples), the grouping (13, (6, 8)) observed predominately in S+ can only be explained by the existence of a situation model which integrates the presence of the structuring element; otherwise, the linear processing would give groupings of the type ((13, 6)…), and this is what was observed in S-. It therefore seems that processing which uses situation models dominates, but when it cannot operate, due to the presence of the structuring term in the instructions, the linear procedural model takes over the role.

Conclusion

The experiment showed the effect of a term which structured the elements of the text on the choice of the solving strategy; this term only came into play by means of the transitory situation model constructed during reading. This result corresponds with Reusser's (1989) Situation Problem Solver (SPS) and, generally, the models which integrate more qualitative representations than the strict formal schema. However, in the absence of a structuring term, the representations are less organised, and it seems that linear processing takes over.

Metacognitive Knowledge in Text Comprehension: Some Issues in Development and Individual Differences

P-E. Eme & C. Haro

University of Poitiers (France)

Elsa.Eme@mshs.univ-poitiers.fr

Metacognitive knowledge in reading may be defined as the knowledge a person possesses about a) the structure and function of written discourse (Gombert, 1990), and b) three broad categories of reading skills: Evaluation, planning, regulation (Paris & Jacobs, 1984). Evaluation refers to the assessment of the task difficulties and one's cognitive abilities. Planning involves the selection of specific strategies to reach specific goals. Regulation concerns monitoring and remediating difficulties during the course of reading. Several authors have pointed out the role of metacognitive knowledge in text comprehension (Garner, 1987; Brown, Armbruster & Baker, 1986; Hacker, 1997). For example, Garner (1987) suggested that in order to read and study expository texts

effectively, students must possess both conceptual knowledge (about the domain or situation dealt with in the text), text-processing strategies (e.g., ongoing summarization and backtracking) and "schema for exposition". Garner defined the latter as a person's knowledge about textual and metatextual signals of the importance or salience of information in a particular genre or type of text. Brown, Armbruster and Baker (1986) reported some differences in metacognition between readers at different ages or reading levels. And Hacker's results showed that metacognitive deficiencies could persist in college students. A unique French study (Ehrlich, Kurtz-Costes, Rémond & Loridant, 1995) showed that the level of metacognitive knowledge could explain a specific proportion of variance in comprehension (multiple regression analysis), even with decoding and motivation variables entered in the regression. However the relationship of metacognitive knowledge and the efficiency of comprehension processes is still unclear. Firstly, studies have focused on children, so that we know very little about knowledge in adults. Second, few studies have attempted to relate qualitatively subjects' comprehension level and their actual knowledge about comprehension: For example, which category(ies) of knowledge are the poor comprehender missing?

The purpose of the present studies was (a) to design a tool in order to assess explicit metacognitive knowledge in both young adults and children and (b) to analyze the development of metaknowledge in children and qualitative individual differences between good and less good comprehenders and (c) to describe knowledge about reading in University students.

Method

A questionnaire (MK23) was designed to collect qualitative data about children and adults' metacognitive knowledge. Four dimensions of metaknowledge were considered:

1) Evaluation, e.g.: "When you're reading, what is difficult for you?"; "Do you always understand what you read?" (5 items);

2) Functions of text features and documents, e.g.: "What is a table of content, what is it for?"; "Do you think that initial sentences in a paragraph or a text are special? Do they have a special importance?" (6 items);

3) Planning, e.g.: "What is the best way to remember what you have read?"; "If you have only 3 minutes to read a 2-page text about which you will have to answer questions later, how do you proceed?" (7 items);

4) Regulation, e.g.: "What do you do when you read a sentence that you don't understand?"; "Do you happen to look back and read again some sentences in a text? If so, why?" (5 items).

Questions aimed at assessing knowledge about these four dimensions were constructed from scratch or based on previous works (Ehrlich et al., 1995; Myers & Paris, 1978).

84 children from grades 3 and 5 (9 and 11 year old) and 66 3rd year psychology students at the University of Poitiers, responded to the MK23 questionaire. Children were interviewed individually and the interviewer transcribed their answers on a scoring sheet. Adults wrote down their answers. Furthermore, children read a 424-word written text for comprehension and answered 24 literal and inferential questions, which provided a comprehension score. Each child was asked to provide an estimate of his/her comprehension score.

Results

Data analysis consisted in scoring all the protocols and identified between 3 and 7 answer categories per item; then the categories were grouped into "simple" answers, "elaborate" answers, and "none/other" answers. The main results are summarized below.

A minority of children demonstrated some explicit knowledge about text features and comprehension activities. Generally, the children didn't provide an accurate evaluation of their comprehension level; they explained their reading difficulties by referring to elementary processes, such as word identification, and not to deeper processes, such as knowledge base or global coherence; they could define text features (paragraphs) and documents (table of contents) but not their function in comprehension; they didn't know many strategies to plan or regulate their activity and they didn't report strategy variation as a function of reading goals. Simple and elaborate answers moderately increased between the ages of 9 and 11 (a significant difference was found only on the "text features" scale). Nevertheless good and poor comprehenders differed in the quality and amount of knowledge, particularly on "planning" and "regulation" subscales.

As expected, elaborate answers increased in adults; in particular, evaluation knowledge seemed to develop during secondary and higher education, as evidenced in University students' rather high scores to these items. In contrast, less than half the undergraduate students provided elaborate answers to questions about planning and regulation activities: Most of them attributed their failure to local features or elementary processes, and overestimated their comprehension level. At last, many adults didn't think about the most useful strategies for solving difficulties.

Discussion

In sum, children's knowledge about reading appeared to be related to their performance in text comprehension. In adults, large individual differences were observed too, which could be related to their academic success or failure. While a great deal of attention has been given to the development of metacognition in children, very little is known on the relationship between meta-knowledge and cognitive performance on comprehension tasks in adults. Yet this issue is becoming critical as many higher education institutions promote learning situations which rely on students' ability to interact with complex information systems. Further studies will assess the possible relationship between students' metaknowledge and their academic success.

References

Brown, A.L., Armbruster, B.B. & Baker, L. (1986). The role of metacognition in reading and studying. In J. Orasanu (Ed.) Reading Comprehension: From Research to Practice (pp. 49-75). Hillsdale, N.J: Erlbaum.

Ehrlich, M.F., Kurtz-Costes, B., Rémond, M., & Loridant, C. (1995). Les différences individuelles dans la compréhension de l'écrit: Facteurs cognitivo-linguistiques et motivationnels. Cahiers d'acquisition et de la pathologie du langage, 13(1), 37-58.

Garner, R. (1987). Strategies for reading and studying expository text. Educational Psychologist, 22, 299-312.

Gombert, J.E. (1990). Le développement métalinguistique. Paris: Presses Universitaires de France.

Hacker, D.J. (1997). Comprehension monitoring of written discourse across early-to-middle adolescence. Reading and Writing: An Interdisciplinary Journal, 9, 207-204.

Myers, M., & Paris, S.G. (1978). Children's metacognitive knowledge about reading. Journal of Educational Psychology, 70(5), 680-690.

Paris, S.G. & Jacobs, J.E. (1984). The benefits of informed instruction for children's reading awareness and comprehension skills. Child Development, 55, 2083-2093.

The Development of Analogical Processes in Learning to Read in French

I. Brun, J. Faure, J. Ecalle & A. Magnan

University of Lyon 2 (France)

Annie.Magnan@univ-lyon2.fr

Several studies have been conducted to test for both assembled and adressed phonologic mediation in reading comprehension (Coltheart, Curtis, Atkins & Haller, 1993). It is clear now that adult skilled reading of frequent words implies two kinds of processes. The adressing refers to a recognition of extracting some visual properties of the stimulus, that will later be used to search the mental lexicon. The assembling refers to the analysis of words' components. The developmental question is how this balance between these two types of processing is achieved. Two alternatives have been proposed. Either the beginning readers develop a sequential reading (letter-by-letter or grapheme-by-grapheme strategy) and later they use bigger units or the beginning readers start reading by focusing on global characteristics of words and later they would focus on detailed information. These two alternatives coincide with the distinction between developmental models. Thus, the stage developmental models (Frith, 1986; Harris & Coltheart, 1986; Morton, 1989) claims that children in the phonological stage first use small units like letters and only later employ larger unit such as morpheme or analogies based on word rimes. The second alternative proposes exactly the reverse. Goswami and Bryant (1990) proposed that the use of analogies based on word rimes appears first and phonological procedure based on grapheme-phoneme correspondence only later. Thus, studies in English language suggest that onset-rime awareness might be connected to children making analogies between the spelling patterns representing those linguistic units in words (Goswami, 1994; Goswami & Mead, 1992; Muter, Snowling & Taylor, 1994; Peterson & Haines, 1992). The first kind of models considers that reading by analogy is characteristic of the last stage of reading acquisition. The second type of models suppose that children start the reading process through an analogical process. More precisely, Goswami and Bryant (1990) think that the children's capacity of analyzing the speech stream into syllables and more specifically into onset and rimes would allow them to discover the existence of analogous and, therefore, to expand their reading skill to new words. Thus, the early stage of reading acquisition distinguish between a small unit approach and a large unit approach. They differ mainly in theirs claims about the size of the linguistic units which are important in the initial stage of reading acquisition. The purpose of this paper is to address the predictions of the " large unit theory " in French. This is approached by using an identification lexical task with pre-literate children and a decision task in the second year of learning.

Experiment 1: Word recognition task in kindergarden

44 pre-literate children of kindergarden (mean age 5;7) were tested to recognize words in a forced-choice task. Two groups (good phonological awareness and poor phonological awareness) were proposed to recognize target words (CVC) with a clue word (analogy condition) or not. Target words and clue words share some orthographic units, either rime (lac-sac), either body (bus-but) either onset+coda (mur-mer). It is shown that only good phonological awareness group significantly recognize in the analogy condition target words which share the rime with clue words, as in Goswami's studies.

Experiment 2: Lexical decision task in second grade

In this lexical-decision task children indicated whether presented letter strings constituted words or not. In this experiment we manipulated type of misspellings (control and analogous pseudo-words), base-word frequency and the reading level. The main question was whether second graders show evidence of analogy. This was investigated by means of a lexical-decision task in which subjects have to decide whether or not individually presented letter strings constitute words. This experiment compared the children's speed and accuracy in lexical decision task of bisyllabic nonsense words. The children were from second grade. A two-groups reading level design was used in this study (Khomsi, 1997). A frequency table (Brulex) was consulted for the selection of twelve bisyllabic words (6 high-and 6 low frequency). The twelve base words were changed into 24 analogous pseudo-words. The stimuli consisted of lists of bisyllabic nonsense words that either shared orthographic units with real words in beginning (garale/garage is a analogous pseudo-word on the beginning APWb) or shared orthographic unit with real words at the end, that is analogous pseudo-word on the end (marage/garage is APWe) and control pseudo-words (pudole). Finally, 36 words were used to serve as fillers. Proportions of correctly identified errors- that is, correct non-word decisions- were calculated for two kinds of analogous pseudo-words and control pseudo-words derived from frequent and infrequent base words.

Results

A 2 (reading level: Good versus poor readers) by 3 (type of misspelling: APWb, APWe, control) and by 2 (base-word frequency: Frequent versus unfrequent) ANOVA on correctly identified errors was performed on subject and item means. Subject's mean decision latencies of correct no-responses to analogous pseudo-words and correct no-responses to control misspellings were also computed. The results of the lexical decision task show that the analogous pseudo-word APWb (garale) were processed faster than the analogous pseudo-words who have kept the final orthographic unit (marage). No frequency effect was found. In good readers, positive evidence was obtained for analogical process because fewer analogous pseudo-words APWb were correctly classified as non-words than control pseudo-words and analogous pseudo-words APWe.

Conclusion

Many aspects of the results of these experiments are apparently consistent with the theoretical viewpoint set out by Goswami (1986, 1993). The results show that the reading by large-unit analogies determines the initial stage of acquisition but a late influence occurs with second graders. Therefore, we assume that analogical procedures are used for both pre-literate children and second graders. In other words, the ability to draw analogies might be thought of as another type of self-teaching mechanism for the beginning readers. In this perspective, the analogical processes correspond to the use, by learner, of his or her own basis of knowledge (phonological and visuo-orthographic) in order to increase his or her lexical knowledge. It is thus concluded that

analogical processes may hightlight some developmental stages and processing strategies in the reading acquisition.

Development of Syntactic Connexion in French Children's Narrative and Expositive Texts

A. Viguié

Institute of Human Sciences, C.N.R.S. (Lyon, France)

Anne.Viguie@ish-lyon.cnrs.fr

Studies of the acquisition of monologues have concentrated essentially on narrative discourse. Few have attempted comparative analyses between different genres. In this paper we will study the acquisition of discourse in two textual types (narrative and expositive), in two modes of production (spoken/written) and two different orders of production (spoken-written/written-spoken) in French children and adults. Our research compares the different linguistic forms used to structure the discourse as a function of these variables, paying particular attention to genre.

The general aim of this study is to access the impact of literacy on the child's developing syntactic competence. Blanche-Benveniste uses the term "first grammar " to refer to the grammar used before schooling and "second grammar " to refer to the grammar encountered in school. Romaine (1984) underscores, the impact of learning to read and write "to successful participation and progress in school" (Romaine, 1984, p. 196). Whereas all speakers acquire the first grammar, some of them will have more difficulties to control the second. " Although all children learn to speak and hear without formal instruction, this is not true for reading and writing (Romaine, 1984, p. 207). For Berman (1996), literacy refers to the capacity " to use decontextualised written texts as a source for deriving information and as a means for conveying information to others (…) Cognitively, literacy involves the ability to use writing not merely for communication, but as a epistemic tool, as a means for analyzing and hence transforming one's own knowledge-base ".

In this present communication we concentrate on synctatic packaging in both narrative and expositive texts. Speakers/writers must organize their text at both the level of coherence or meaning, as well as at the level of cohesion, or linguistic structure. Our study examines clause combining from both a syntactic and a semantic perspective. We use the notion of t-units as defined by Biber (1998), in accordance with Hunt (1965), as an independent clause with all its dependent clauses. The level of t-units provides e measure the syntactic complexity. From a developmental perspective we examine both the size of t-units and their internal structure across spoken and written texts and narrative and expository texts. In addition we examine the semantic relations between the clauses, which make up the t-unit.

Method

The work presented is part of larger international project undertaken in collaboration with Israel, Netherlands, Spain, Sweden and United-States. Like the other countries, the French team has included four age groups: three for children and one for adults. The age of the child groups are 9/10-, 12/13- and 15/16-years-old. The mothers of all the children have a baccalauréat. The adults have a level of education of at least 5 years of study after the baccalauréat. Each age group consists of forty subjects. Subjects produce one narrative and one expository text, each text in a spoken and a written mode. Thus there are four texts for each subject, the data collection is organized in two separated parts: session one and session two. Each subject takes part in the two sessions. In addition, the order variable is controlled resulting in four orders, shown in table 1.

Table 1: Order of production

ORDER A: Session 1: Narration Spoken/Narration Written

                    Session 2: Expositive Spoken/Expositive Written

ORDER B: Session 1: Narration Written/Narration Spoken

                    Session 2: Expositive Written /Expositive Spoken

ORDER C: Session 1: Expositive Spoken /Expositive Written

                Session 2: Narration Spoken /Narration Written

       ORDER D: Session 1: Expositive Written /Expositive Spoken

                    Session 2: Narration Written /Narration Spoken

Hypotheses: (H1) The number of clauses per t-units increases with age; (H2) The inventory of clause type changes as a function of age, text type and modality of production; (H3) The variable text genre will prove to play a stronger role than modality of production.

Results

The results confirms the first two hypotheses: The number of clauses per t-unit increases with age. The inventory of subordination types increases with age. Text genre is as important if not more important than modality as a variable. However, in terms of semantic relationships established between successive clauses, the developmental progress is much more difficult to interpret.

Phonological Recoding in Deaf Children

L. Paire-Ficout, S. Colin & A. Magnan

University of Lyon 2 (France)

pairefi@univ-lyon2.fr

There is now substantial evidence from studies of reading acquisition to suggest that phonological skills are related to the development of young children's reading ability in alaphabetic orthographies (Bradley & Bryant, 1985; Stuart & Coltheart, 1988; Wagner, 1988; Wagner, Torgeson & Rashotte, 1994). In French, the syllable is a phonologically and orthographically well-defined unit. It proved of particular importance in speech perception of words (Melhler, Dommergues, Frauenfelder, & Segui, 1981) and in speech production (Ferrand, Segui, & Grainger, 1996). So the role of syllables as phonological units in hearing and in deaf children has still not raised a lot of research in French. Children with severe/profound prelingually deafness encounter considerable difficulty in achieving fluent reading skills. Generally, it has been hypothesised that the relatively poor reading skills of mostly deaf individuals results from deficiencies in phonological processing. Indeed, the lack of accurate and specified phonological representation results from the lack of adequately developed spoken language. For example, Conrad (1972) found that hearing subjects confused words which sounded similar, whereas deaf subjects confused similar spelled words. More recently, Merrils, Underwood and Wood (1994) suggested that deaf readers relied more on visual word features than hearing readers did. In the same way, Waters and Doehring (1990), Harris and Beech (1997) did not provided evidence for deaf subjects sensitivity to the regularity of pseudohomophones. Moreover, Campbell and Wright (1990) used a paired-association task to show that unlike hearing children, deaf children did not use picture-name rhyme spontaneously as a cue to recall pairs. In contrast to these results, some authors found that deaf children are able to develop and use a phonological code in various tasks. For example, deaf subjects were able to read pseudo-words just like hearing subjects (Leybaert & Alegria, 1993). Evidence showed that pseudo-words with a complexe phonology were more difficult to read than those with a simple phonology. Deaf subjects were also more accurate in reading regular than irregular words. Transler, Leybaert and Gombert have shown that deaf children like hearing children use syllables as phonological units in print processing situations. Children were asked to copy written words and pseudo-words and measured the number of glances at the item, copying duration and locus of the first segmentation. Like hearing children, deaf children used syllables as copyong units. They only differed in a condition in wich orthographic and phonological criteria were differentiated. We also collected data providing evidence that profoundly deaf children (8;6 years) like hearing (see Colé, Magnan, & Grainger, in press; Magnan & Colé, 1999) were sensitive to the syllable compatibility between target and words, but only in case of short and familiar words (Paire-Ficout & Magnan, accepted).

The aim of this present study is to compare deaf children who have been exposed to Cued Speech (CS) and deaf children who have not been exposed to CS. In CS, speech/reading is combined with manual cues which disambiguate it. The central hypothesis is that young deaf children with greater exposure of Cued Speech would be more sensitive to syllabic congruency. In order to test this hypothesis, we used a visual version of the monitoring technique employed in studies of speech perception (Mehler et al., 1981). Subjects were asked to detect a letter string target presented visually at the initial position of words. The target structure was either a consonant-vowel (CV) or consonant-vowel- consonant (CVC) structure and either did or did not correspond to the structure of the first syllabic part of the target bearing words. Thus, for example, children were asked to detect a letter string BA (CV target) in BALLON (CV word) or in BALCON (CVC word) and the letter string BAL (CVC target) in BALLON and BALCON. This task can be performed phonologically or visually and if the beginning readers process written language on visual/orthographic grounds only, the detection times should only be function of the size of the target sequence so that detection times should be faster for CV targets (2 letters long) than for CVC targets (3 letters long) whatever the syllabic congruency is. In contrast, if the task is performed phonologically, the detection times should be function of the syllabic congruency between the structure of the target sequence and the one of the first part of the test word. As an example, detection times should be faster when detecting SO in SOLEIL (sun) than SOL in SOLEIL and detecting SOL in SOLDAT (soldier) should be faster than detecting SO in SOLDAT. In the latter case, an interaction between the size of the

target sequence and the syllabic structure of the test word is expected. The first data seem indicate that Cued Speech allows deaf children to be more sensitive to the syllable congruency and to detect the congruency more rapidly.

The Influence of the Phonological and Orthographic Characteristics of French During Reading Acquisition

J. B. Aimar* & A. Magnan**

***University of Nice-Sophia Antipolis (France)**

****University of Lyon 2 (France)**

aimar@unice.fr

Numerous findings show that phonological processing skills are fundamental in the acquisition of literacy (Goswami & Bryant, 1990; Liberman & Shankweiler, 1985). According to Wagner and Torgesen (1987) phonological processing is made of three sets of distinctive component skills: Phonological awareness, phonetic recoding in working memory and phonological recoding in lexical acces. Phonological recoding involves the extraction of lexical characteristics and pronunciation codes from written information. Thus, in this phase, the beginning reader uses the relations between graphic units and phonological units in reading. Most current theoretical models of reading acquisition suggests that reading acquisition rapidly allows syllable-sized units to be accessed from print (Sprenger-Charolles & Siegel, 1997), and that this type of coding continues to influence how adult readers process low-frequency words (Colé, Magnan & Grainger, 2000; Ferrand, Segui & Grainger, 1996). The syllabic effects obtained with good readers after the first year of schooling (Colé & Magnan, 1997; Magnan & Colé, 1999) may be specific to French and more precisely to the kind of orthographic system. French is a language where the syllable is likely to be a critical unit in the processing of speech. We hypothesize that French children at the start of reading acquisition use their syllabically structured spoken-words system. French, although not perfectly regular in the transcription from phonemes to graphemes is more regular than the other way around. Thus, it is a much more transparent language than English and has a much more predictable relation between graphemes and phonemes. The main purpose of this work is to show that this syllable-based procedure is influenced both by the phonological and orthographic characteristics of French. More precisely, during reading acquisition this syllabic process might be affected when the orthographic transcription of syllables is difficult i.e contains a consonant cluster in the onset (CCV). So we used words with first syllables that possess a consonant cluster in the onset (CCV) or a rime with a coda (CVC). In French, performance on closed syllables with a coda (CVC) can be expected to be better than performance on open syllables (CCV). More precisely, inversions of letter sequences or omissions of one letter, often pointed by teachers, ("cravate" (tie) is read " carvate " or " cavate ") are considered like visual errors. We think that these sequential errors constitute a partial use of grapheme-phoneme correspondences (Bastien-Toniazzo, Magnan, Bouchafa, 1996, 1998; Sprenger-Charolles & Siegel, 1996). The basic structure of the syllable is a simple onset and a rime without coda (Clement & Keyser, 1983). This single consonant and vowel (CV) exists in all languages and leads to superior performance in the early acquisition of reading.

Method

Subjects. A group of 20 children (mean age in december was 6 years 7 months) was tested in the months of December and June during the first grade which is the first year of reading instruction in France. Using a reading test (Khomsi, 1990), we distinghish poor and normal readers. The subjects were selected according to the following criteria. All were average with regard to cognitive development. All were native French speakers from the middle social classes, who were attending

their grade at the regular age. They also had normal (or corrected-to-normal) vision and normal hearing. The children came from 2 classrooms from a school in suburbs of Lyon. The teachers used the same " mixed " method. This method combined a " global " approach (using key words and short texts) and an " analytical " approach (focusing on simple vowels and consonants in syllables and in words).

Material and Procedure. Children were asked to detect a letter string target presented visually at the initial position of word presented below. The target was either a consonant-consonant-vowel (CCV) or a consonant-vowel-consonant (CVC) structure and either did or did not correspond to the first Syllabic part of the target bearing word. For example, children had to detect CAR (CVC target) in CARNET (CVC word) or in CRAPAUD (CCV word) and CRA (CCV target) in the same words. The same word was associated with two different experimental conditions. Four experimental situations were proposed to children. The task was performed with familiar words. Familiarity was assessed by a self made test. The word length was controlled (6/7 letters). The same words were used for two sessions. The children were tested individually in a 15-20 minutes session. Sequences were presented in the center of an computer monitor as follows. A fixation point was presented during 500 milliseconds. After a 500 milliseconds blank interval, the target was presented on the screen during 500 milliseconds then the test-word appeared directly. The test word remained on the screen until the child's response. He or she had to decide as quickly and as accurately as possible whether or not the target occurred At the initial position of the test-word. The child had to press one button If the target and test word corresponded and another one if they did not. The "yes" response was given with the preferred hand. The experimental session began with 5 practice trials. The sequences order was randomized. We think that the first good readers after one year of schooling will be sensitive to the congruence of the target and the word. However, no such matching is expected with consonant clusters at the beginning of learning. Our hypothesis is that the processing varies as function of the complexity of syllables. More precisely, we think that CVC words would produce significantly fewer errors than CCV words. The consonant clusters could involve a delay in the processing. The congruence is expected for CCV words and CVC words in the session 2 only for the good readers.

Results

The results suggest that good readers activate orthographic and phonological informations more efficiently than poor readers. Our results agree with the recent ones of Booth, Perfetti and Mac Whinney (1999), showing that becoming a skilled reader involves the development of fast and automatic orthographic and phonological processes. We think that poor readers have less efficient orthographic and phonological processes that may result from the imprecise and sparse nature of their representations (Perfetti, 1992).

Select-a-Kibitzer: A Multi-Agent Architecture for Giving Feedback on Student Compositions

P. Wiemer-Hastings* & A. C. Graesser**

*University of Edinburgh (Scotland)

**The University of Memphis (U.S.A.)

peterwh@cogsci.ed.ac.uk

Language teachers are faced with a difficult dilemma. They know that students benefit from practice, and from getting feedback on their compositions. But the more in-depth the feedback is, the more time it takes to produce the feedback. With finite amounts of grading time, language teachers are thus faced with a tradeoff: Either assign less writing or provide quicker, less-detailed feedback. This paper describes a computer tool which uses a variety of language processing technologies to address this dilemma in a novel way. The tool is called Select-a-Kibitzer (hereafter SaK). SaK is based heavily on theories of language composition that describe the process of writing as one of attempting to simultaneously solve multiple, potentially conflicting constraints (Flower, 1994). These constraints come from a wide range of sources. Some constraints depend on the student's knowledge (or lack of knowledge) of their language. Spelling, grammar, and word sense selection are difficult for many students. Other constraints are conceptual. Students have difficulty reasoning about what they know about an assigned topic and what they can relate in a composition. Other constraints are socio-linguistic in nature. There are commonly differences between the teacher's language, and the language the students naturally speak. The existence of so many different constraints makes the composition process very difficult. What makes it more difficult is that students normally do not even know explicitly what the different constraints are. Thus, they are unable to reason about how to reconcile them. SaK attempts to make the student's task easier by making the constraints explicit. It does this by using different computer agents to represent different constraints.

The knowledge which these agents use is called from a variety of sources. Lexical knowledge is extracted from the WordNet lexicographer's database (Fellbaum, 1998). This knowledge source helps determine not just the spelling and sense of a word, but also related words which are more or less specific (hyponyms and hypernyms). Knowledge about the syntactic structure of the student's text is extracted by parsing the text with a readily-available natural language parser (Allen, 1995). SaK process WordNet definitions into a form in which they can be used as an online lexicon for the Allen parser. The parser provides a parse tree for each of the student's sentences.

The sources mentioned above provide important information, but it is also information that is fairly easy for teachers to give feedback on. The most important source of information for SaK is that about the meaning of the text. Latent Semantic Analysis (LSA: Foltz, Kintsch, & Landauer, 1998; Landauer & Dumais, 1997) is a statistical technique which gives impressive results in computing the semantic similarity between texts. By themselves, semantic similarity judgments between texts can be a valuable tool to evaluate text meanings. Given a set of "target essays" for a particular assignment that cover a range of relevant topics, one can determine with LSA which of these essays the student's composition is most similar to, and thereby determine the overall topic of the essay. This technique can also be used to suggest additional material that the student might want to address, and determine if the student's composition is completely irrelevant to the assignment.

The potential uses of LSA go far simple semantic judgments, however. By comparing the LSA scores of consecutive sentences in a composition, you get a measure for how coherent the overall composition is. You also can tell where breaks between different subtopics are. LSA can also be used to provide an automatic summary of a text, by finding the breaks between subtopics, and then taking the most central sentence from each (Wiemer-Hastings & Graesser, in press).

The individual agents of SaK are implemented using the Soar problem-solving architecture (Laird & Rosenbloom, 1996), which allows us to express the agents knowledge in the form of condition-action rules. SaK places the information from each of the linguistic information sources into Soar's working memory. SaK's agent rules use this information along with contextual knowledge about the student and the particular assignment to decide what type of feedback should be given to the student about his or her composition. SaK then displays that feedback to the student. The key is that, unlike a standard computer style checker, the feedback does not come from a single, "all-

knowing" evaluator, but is associated with different characters. Each different agent has an iconic picture (for example, a cowboy, dog, or William Shakespeare). Other studies show that information presented by more-or-less lifelike agents is more readily understood and remembered by students (for example, Lester et al, 1997). Each agent comments on a single aspect of the composition. Some of the comments are negative, pointing out errors or misconceptions by the student. Others are positive, praising some aspect of the student's composition. By recognizing the correspondence between the agents and the types of feedback, the student can recognize that there are different aspects of the writing process which must be taken into account. This knowledge will allow them to think explicitly about what they can do to improve their compositions.

The current status of SaK is that it is has all the pieces in place to collect the linguistic knowledge, to fire the rules which invoke different feedback items, and to present them graphically to the student. What SaK currently lacks is the principle pedagogical knowledge concerning what types of feedback are most appropriate at what stages of the composition process. The authors welcome ideas concerning that aspect of the system.

References:

Allen, J. (1995). Natural Language Understanding. Addison-Wesley, Reading, Mass.

Fellbaum, C. (1998). WordNet: An electronic lexical database. MIT Press, Cambridge, MA.

Flower, L. (1994). The construction of negotiated meaning: A social cognitive theory of writing. Southern Illinois University Press, Carbondale, IL.

Foltz, P. W., Kintsch, W., & Landauer, T. K. (1998). The measurement of textual coherence with Latent Semantic Analysis. Discourse Processes, 25, 285-308.

Laird, J., & Rosenbloom, P. (1996). The evolution of the Soar cognitive architecture. In Steier, D. M., & Mitchell, T. M. (Eds.), Mind Matters: A tribute to Allen Newell (pp. 1-50). Erlbaum, Mahwah, NJ.

Landauer, T., & Dumais, S. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of acquisition, induction, and representation of knowledge. Psychological Review, 104, 211-240.

Lester, J., Voerman, J., Towns, S., & Callaway, C. (1997). Cosmo: A Life-like Animated Pedagogical Agent with Deictic Believability. In Working Notes of the IJCAI '97 Workshop on Animated Interface Agents: Making Them Intelligent or Special Issue of the Applied Artificial Intelligence Journal, pp. 61-69. London. Taylor and Francis.

Wiemer-Hastings, P., & Graesser, A. (in press). Select-a-Kibitzer: A computer tool that gives meaningful feedback on student compositions. To appear in Special Issue of Interactive Learning Environments, Joseph Psotka (Ed.).

ETAT: A Tool to Analyse Expository Text Coherence

E. Vidal-Abarca*, H. Reyes*, R. Gilabert*, J. Calpe*, E. Soria* & A. C. Graesser**

*University of Valencia (Spain)

**The University of Memphis (U.S.A.)

Eduardo.Vidal-abarca@uv.es

Qualitative methods that analyze the coherence of expository texts are not only time consuming, but also present challenges in collecting data on coding reliability. We describe software which analyses expository texts more rapidly and produces a notable level of objectivity. ETAT (Expository Text Analysis Tool) analyzes the coherence of expository texts. ETAT adopts a symbolic representational system, known as "Conceptual Graph Structures" created by Graesser (1981). Graesser´s system represents knowledge structures, but it might also be looked upon as a system that represents text information. The information is represented as a network of labelled statement nodes that are interrelated by labelled, directed arcs. The main difference between ETAT and the original CGS procedure is that the latter represents statement nodes coming from the text and also from the reader's inferential activity, whereas ETAT only represents the information in the text. CGS representational system has directly or indirectly been validated as being psychologically plausible in numerous empirical studies (Graesser & Hemphill, 1991; Zwaan, Langston, & Graesser, 1995).

Analysis Procedure

ETAT aids the analysis of text information according to the three steps (node segmentation, node classification, and node connection) as described for the construction of the CGS. However, for each step, some changes have been introduced to simplify the analysis and adapt the tool to any expository text. In the first step, the text is divided into nodes defined as complete sentences. They are grouped into blocks, with each block being formed by a maximum of 14 nodes, approximately 2-3 paragraphs of text. In the second step, a category is assigned to each node, based on its conceptual content. ETAT classifies the nodes in one of three classes: State (S), Event (E), and Goal (G). For the third step ETAT recognizes six possible conceptual relationships between nodes: Initiate, Reason, Outcome, Cause, Description, and Example. The first four categories were taken directly from Graesser Wiemer-Hastings and Wiemer-Hastings (in press), whereas the last two groups diverse semantic relationships formulated by Graesser et al. (in press), Meyer (1985), and Bovair and Kieras (1985). Each relationship has a composition rule which specifies the node categories that can be linked by a particular type of relation. For example, only an event (E) or a state (S) can initiate a goal (G). All relationships can be either explicit or implicit. Once the relationships between all pairs of nodes and each one of the blocks have been established, ETAT automatically selects the nodes with the greatest number of relations within each block and are then grouped into new blocks. The analyst once again compares the relationships between nodes in the new blocks. The process of selection and regrouping of nodes, and interrelationship testing is repeated until one unique block is obtained which then completes the analysis.

Besides the six previously described relationships with their conceptual characteristics, ETAT permits the automatic establishment of another non-directional, optional relationship between pairs of nodes: Keyword Overlap (K). It is established between nodes that share terms with the same etymological root. For example, in a text about the Russian Revolutions, if two nodes share terms with the etymological root "Russ-" (e.g., Russia, Russian), ETAT automatically establishes the overlap relationship between them. The analyst decides which keywords of the text should be considered to establish this relationship, normally those that can be found in the title or headings within the text, generally no more than 2-3 words. As a product of the analysis, ETAT produces three results: Graphs, Isolated-nodes, and Statistics. A graph represents the interrelationships

between the nodes. It is possible to generate a graph with only the nodes of a particular block, of a particular phase or all of the nodes of the text. With the Isolated-nodes option ETAT provides information about nodes or groups of nodes that are not interrelated to the rest of the nodes. ETAT also provides a series of statistics with respect to the interrelationships found in the analyzed text: (a) number of interrelationships between nodes (e.g., Initiate, Outcome, or Total), (b) number of explicit or implicit relationships, (c) means and and standard deviations of the number of interrelationships per node, or (d) proportion of explicit interrelationships over the total number.

Technical Description of ETAT

The program was developed in Java with a compiler Jbuilder. It can be used on a PC, a MAC, and any other system in which a virtual Java engine has been implemented. The program functions correctly with a 486 processor, a minimum of 16 Mb of RAM memory, and Java support. ETAT occuppies 30 Mb of hard disk space. ETAT has the normal features of a Windows environment, such as multiple menus and menu options, selection of options with a mouse and point-and-click facility, and a keyboard for alphanumeric input. The interface has been designed to have rapid access to all information which is produced in the text analysis process.

Uses of ETAT

ETAT can be used to analyze breaks of local and global coherence in studies of text revision. Researchers have used methods that are either time consuming or partially intuitive (Britton & Gulgoz, 1991; Beck, McKeown, Sinatra, & Loxterman 1991; Vidal-Abarca et al., 2000). ETAT will reduce the time to perform these analyses. Investigators in text processing can also take advantage of ETAT. It could help them in objectifying and quantifying structural characteristics that produce differences among texts in text processing research. ETAT could also make the task of textbook publishers and editors of expository text in general more objective and less time consuming.

References

Beck, I. L., McKeown, M. G., Sinatra, G. M., & Loxterman, J. A. (1991). Revising social studies text from a text-processing perspective: evidence of improved comprensibility. Reading Research Quarterly, 26, 251-275.

Bovair, S., & Kieras, D. E. (1985). A guide to propositional analysis for research on technical prose. In B. K. Britton & J. B. Black (Eds.), Understanding expository text (pp. 315-362). Hillsdale, NJ: Lawrence Erlbaum Associates.

Britton, B. K., & Gülgöz, S. (1991). Using Kinstch's computational model to improve instructional text.: Effects of repairing inference calls on recall and cognitive structures. Journal of Educational Psychology, 83, 329-345.

Graesser A. C. & Hemphill, D. (1991). Question answering in the context of scientifc mechanism. Journal of Memory and Language, 30, 186-209.

Graesser, A. C. (1981). Prose comprehension beyond the word. New York: Springer-Verlag.

Graesser, A. C., Wiemer-Hastings, P., & Wiemer-Hastings, K. (in press). Constructing inferences and relations during text comprehension. In T. Sanders, J. Schilperoord and W. Spooren (Eds.), Cognitive approaches to text coherence. Linguistic and psycholinguistic aspects of discourse representation. Benjamins: Amsterdam/Philadelphia.

Meyer B. J. F. (1985). Prose Analysis: Purposes, Procedures and Problems. In B. K. Britton & J. B. Black (Eds.), Understanding expository text (pp. 11-64). Hillsdale, NJ: Lawrence Erlbaum Associates.

Vidal-Abarca, E., Martinez, G., &, Gilabert, R. (2000). Two procedures to improve instructional text: Effects on memory and learning. Journal of Educational Psychology, 92, 1, 1-10.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. Psychological Science, 6(5), 292-297.

Insights into the Interactive Process of a Computerized Tutor:A Focus Group Study

R. DiPaolo & S. Rajan

The University of Memphis (U.S.A.)

rachel@thoughtware.com

AutoTutor is a computer simulation involving an animated pedagogical agent. Animated pedagogical agents are a synthesis of computer-animated characters and interactive learning environments. AutoTutor differs from other pedagogical agents, because he imitates the gestures and dialogue of a human tutor within the context of a tutoring session. AutoTutor attempts to generate the verbal and nonverbal characteristics of a human tutor, while guiding students through the knowledge domain of basic computer literacy. The goal of trying to emulate a human tutor is theoretically grounded in the literature on the effectiveness of tutors. Many studies have documented the claim that one-on-one instruction, such as tutoring, is superior to traditional classroom instruction (Bloom, 1984; Cohen, Kulik, & Kulik, 1982; Graesser & Person, 1994; Hume, Michael, Rovick, & Evens, 1993; Merrill, Reiser, Ranney, & Trafton, 1992; Mohan, 1972; Trismen, 1981). The effect size of the advantage of a tutoring style over classroom ranges from .4 to 2.3 standard deviation units (Bloom, 1984; Cohen, et al., 1982; Mohan, 1972). Therefore, great potential lies in an animated agent that could take the place of a human tutor, allowing convenience and accessibility for students. Four focus groups were conducted in order to collect rich, descriptive, information about the computer-human interaction with AutoTutor. The focus groups should provide qualitative data on students' perspectives regarding empirical questions that have not been answered in the literature. We had two general goals in conducting focus groups. One was to identify potential logistical problems with the current system and the other was to gather novel information regarding the computer tutor - student interaction. More specifically, the goals were accomplished by the following: (a) listening and learning from participants' ideas about AutoTutor, (b) using that information to offer suggestions about adjustments and additions to the current program, and (c) gaining information about student receptiveness to this type of learning tool. All these goals were accomplished through the focus group methodology.

We conducted four focus groups consisting of students from a basic computer literacy course at the University of Memphis. The participants were already familiar with the content that AutoTutor covers during the tutoring session because the material is taken directly from the course. However, in order to have the participants engage in a lively discussion of AutoTutor, they needed a recent experience with the system. Therefore, before the focus group discussion, the participants individually interacted with AutoTutor. The groups consisted of four to eight participants. The groups were medium sized, allowing each person an opportunity to express his or her thoughts and opinions within the allotted time. The moderator played a vital role in maintaining the flow of discussion and making sure that everyone had an opportunity to talk. The focus groups followed a particular sequence: (a) opening with a brief discussion, (b) interacting individually with AutoTutor, (c) moving to another room, (d) taking a short break with refreshments, and (e) participating in a focus group discussion about the participants' experiences with AutoTutor.

The focus groups, moderately structured, incorporated a funnel design for a questioning route. The funnel design involves opening with two or more broader questions and as the discussion progresses, narrowing to more specific questions. There were a total of twelve questions asked in the focus groups. The twelve questions fell into the following categories: one opening question, one introductory question, two transition questions, four key questions, two key-key questions, follow-up question (the number of follow-up questions varied), one summary question, and one closing question. However, during the analysis, we found that some of the questions fell into different categories from what we had anticipated. Question three, a transition question, provided extensive and valuable information. Therefore, it played the role of a key-key question. As stated earlier, the question format adhered to a funnel-design questioning route. The focus groups accomplished their intended goals, which were mentioned earlier. After listening to the tapes, we highlighted the key findings from each question. We classified the findings into various topics that we based on the questions' contents and the students' responses. We organized all our notes according to the topics, forming four main categories. These notes were a combination of the written notes, taken by an assistant, as well as, the notes taken from the tapes for each group. Additionally, any unexpected topics or new ideas were identified and included. The final set of notes were summarized and interpreted.

Within and across the four focus groups several points consistently reoccurred: (a) lack of cartoon animation, (b) insufficient feedback, and (c) limited communication. The lack of animation produced by AutoTutor affected the interest and learning of the participants. The participants often commented on his minimal use of facial expressions and when he did produce them, the duration of the expression was too short. When the participants did notice the facial expressions, most found them helpful in understanding AutoTutor's overall response to their contribution. However, their awareness of the facial expressions rarely occurred. Several points were consistently mentioned within and across the four focus groups; however, one of the most prevalent points related to dialogue between AutoTutor and the participant. In general, the participants commented that communication was limited with the system, verbal and nonverbal. This limitation affected the feedback and the pedagogical effectiveness of AutoTutor. The limited verbal communication was derived from two sources, difficulty in understanding the computerized voice and lack of AutoTutor's ability to respond.

AutoTutor tries to mimic not only the verbal aspects of a human tutor, but the nonverbal as well. Facial expressions were the main source of nonverbal feedback for this phase of the project. One unexpected finding was that the facial expressions often went unnoticed. The ones that did notice found the expressions helpful in understanding AutoTutor's overall response to their contribution. However, these participants often commented on his minimal use of facial expressions, and when he did produce them, the duration of the expression was too short. As a result of the suggestions that came out of the focus groups, we have many feasible ideas on how to improve the deficiencies. The limited communication affected the feedback in that the participants rarely knew if the contribution they gave was right or wrong due to either (a) AutoTutor ignoring what they typed, (b) AutoTutor not recognizing what they typed, or (c) AutoTutor giving a vague response

(e.g., "I see," "okay", "well"). There are already many solutions to these problems that are currently being designed and implemented.

From the data collected from the focus groups we have been able to diagnose problems much more accurately and provide recommendations for future changes in the system to make AutoTutor more conversationally smooth as well as a more pedagogical system. Overall, the participants were excited about the idea of a system such as AutoTutor supplementing the course as a learning and review tool.

Short Responses in Human and Intelligent Tutoring Systems

S. Rajan, D. Harter & A. C. Graesser

The University of Memphis (U.S.A.)

srajan@memphis.edu

Short responses are an important type of discourse feedback. According to Clark (1996), the purpose of short responses is to help coordinate the exchange of information within the dialogue. The presence of short responses involves not only verbal utterances but also different kinds of non-verbal utterances that form an essential part of the communicative process. Interactional discourse functions are responsible for creating and maintaining an open channel of communication between the speech participants, whereas propositional functions shape the actual content (Cassell, Bickmore, Billinghurst, Campbell, Chang, Vilhjalmsson, & Yan, 1999). Since the presence or absence of feedback affects the mutual behavior of the discourse participants (Cassell et al., 1994; Clark, 1996), exchange and receipt of information within tutoring sessions is presumably facilitated by means of these short responses. The actions of the listener vary with the actions of the speaker, so conversation is highly reactive between the speech participants (Duncan, 1974). A single short response might perform different functions such as 'I hear', 'I acknowledge what you say', 'I understand', 'I agree', or 'I am paying attention' (Clark, 1996; Nofsinger, 1991) that might help to differentiate between active and passive short responses (Drummond & Hopper, 1993; Jefferson, 1984/1993).

This thesis determines the frequency of various types of short responses as a function of linguistic locations (speech act boundary, noun phrase boundary and non-boundary) and the speaker role (student versus tutor). The present research performs a systematic analysis on an existing corpus of 27 human-to-human tutoring sessions (Graesser & Person, 1994; Graesser et al., 1995) with regard to short responses. The participants and conversations that were sampled from the corpus of tutoring sessions, consisted of undergraduate psychology students being tutored on various topics in a Research Methods course by three psychology graduate majors at the University of Memphis. The three tutors had received grade A in their undergraduate as well as their graduate Research Methods class. The tutors had prior tutoring experience, but received no formal training. From each of these sessions, 90-second segments were randomly selected for the three tutors who were tutoring five different undergraduate students each.

The analysis of short responses in the human tutoring sessions was divided into three categories: (a) yeah/yes, ok, right, (b) uh-huh, and (c) head nods (down-up, down-up movement) based on previous research (Drummond & Hopper, 1993; Duncan 1972; Fox, 1993; Jefferson, 1984/1993). The occurrence of each short response was identified as occurring in one of three linguistic locations: At a speech act boundary, at a noun phrase boundary and at any other temporal location (non-boundary). For the purpose of this study, we combined the occurrence of short responses that occurred at the end of a speech act and at the end of a noun phrase to indicate feedback at a boundary, versus feedback that occurred at non-boundaries. The data collected on each type of

short response shows the frequency of occurrence and the location where a particular short response is used. The summarized data is useful in identifying trends in the frequency of occurrence of various categories of short responses across various linguistic locations in real tutoring discourse.

The results indicate that verbal feedback mainly occurred at junctures and discourse boundaries rather than within phrases and other non-boundaries. Category (a) which consisted of utterances such as "yeah/yes, ok, right" occurred 0.370 times per minute from students and 1.665 times per minute from tutors at boundaries; this compares to 0.022 times per minute from students and 0.088 times per minute from tutors at non-boundaries. The short response "uh-huh" was observed at almost all the linguistic locations. The data suggest that this form of short response occurred more frequently at the boundaries when compared to non-boundaries. This occurred 0.666 times per minute from students and 0.132 times per minute from tutors at boundaries, compared to none per minute from students and 0.066 times per minute from tutors at non-boundaries. This establishes the fact that verbal utterances tend to be less frequently used within a phrase to allow for continuity in flow of conversation.

We also saw that head nods and "uh-huh" play a significant role in regulating the interactive process and turn-taking functions. Nods occur more frequently within a phrase but are also seen at junctures and discourse boundaries as well. Head nods occurred 3.443 times per minute from students and 0.286 times per minute from tutors at boundaries, compared to 5.888 times per minute from students and 0.333 times per minute from tutors at non-boundaries. This demonstrates that the students are more likely to give non-verbal feedback that is unobtrusive and involves no turn-taking functions.

The significance and meaning of short responses changes very often during a conversation. These short responses may demonstrate different functions such as paying attention, listening, or understanding as mentioned earlier (Clark, 1996; Nofsinger, 1991) depending on the emphasis of movement, speed and intonation of the short response. Short responses might also perform the role of continuers, acknowledgers, encouragement, pumping for more information, etc. This allows for a smoother flow of conversation while still keeping the interactive process alive.

Short responses are an integral part of tutoring sessions, yet they have not been implemented as effectively in intelligent tutoring systems. Turn-taking is a necessary element of any conversational system (Sacks, Schelgoff, & Jefferson, 1974). Absence of indicators of turn-taking could result in incomplete information being transmitted and the disruption of conversation (Thurisson & Cassell, 1996). All of the above data form the basis for further developing an intelligent tutoring system, AutoTutor (Wiemer-Hastings, Graesser & Harter, 1998; Graesser, Wiemer-Hastings, Wiemer-Hastings, Kreuz, and TRG, 2000) capable of giving various short responses dynamically and consistent with human-like behavior. We have incorporated and implemented different types of dialogue moves, including short responses, in our current version of AutoTutor. Because we consider short responses as an important feature of realistic human-to-human discourse, we are constantly trying to incorporate more features in the intelligent tutor relating to envelope feedback (Cassell & ThÛrisson, 1999) and paying adequate attention to facial expressions as well.

In our first version of AutoTutor, we have implemented some simple types of verbal and non-verbal short responses in accordance with the data from human-to-human tutoring sessions. This feedback is given in immediate response to a contribution from the student. These simple feedback mechanisms communicate how AutoTutor is reacting to the utterance just produced by the student (e.g. positive, negative, neutral, etc.). Absence of short responses is a serious concern not only in the emulation of a human tutoring situation but also if any learning outcomes are to be gained from working with AutoTutor. These newer mechanisms could be based on the behavior observed in the natural human-to-human tutoring discourse that has been reported here. All attempts are being made to mimic human performance and eventually have an intelligent tutor capable of giving short responses dynamically, which would keep the student's interest alive and lead to valuable learning gains.

Verb-based Classification of Abstract Concepts

K. Wiemer-Hastings & A. C. Graesser

The University of Memphis (U.S.A.)

kwiemer@LATTE.MEMPHIS.EDU

Traditional concept classification models (e.g., Collins & Loftus, 1975) are based on concepts attributes. Abstract concepts do not readily evoke "attributes" in free generation tasks (e.g., Graesser & Clark, 1985; Markman & Gentner, 1993); therefore, abstract concepts are probably classified based on different aspects. Abstract concepts are not perceivable, but they often relate to their contexts of use. One can argue that linguistic context of abstract nouns is important for their acquisition, representation and comprehension (Quine, 1960; Schwanenflugel, 1991). Children know a lot of their language by the time they acquire abstract nouns. Abstract concepts are used in language, but they can remain unnoticed in a situation. Linguistic description can help to make such concepts in a situation salient to a child. It should accordingly be possible to identify context elements that provide a basis for classifying abstract concepts. Predicates are used in the linguistic contexts of the classified objects. Therefore, abstract concept classification by predicates is one option. However, predicates do not distinguish well within ontological classes, and to some extent not even between (Keil, 1979). The present study explores the discriminative power of another context element of abstract concepts, namely verbs, with a text corpus analysis. Such analyses are commonly used to determine the statistical distribution of language features in naturally occurring language (Boguraev & Pustejovsky, 1996). Verbs link together objects and propositions, and may thus reflect complex representations relevant to abstract concepts. This study tests if verbs that co-occur with abstract nouns provide a basis for a more fine-grained classification of abstract concepts, and if the similarity of abstract concepts is related to their verb contexts, similarity.

Characteristics of noun concepts put constraints on the verbs that can be used with them, and vice versa. Abstract nouns that can co-occur with particular verbs should therefore share semantic aspects. In principle, different concepts can co-occur with the same verbs. For example, one can have and listen to a visitor or a frog. However, further verbs, such as jump versus walk, make the difference obvious. The picture is analogous to fuzzy boundaries in prototype structures of concrete objects, with verbs playing the role of object features. Some verbs occur with many different concepts (that belong to different categories), but different concepts will co-occur with some verbs that can only be used with one and not the other concepts. As long as the text sample is large enough, it will likely contain distinctive verbs, so that 1) the co-occurrence frequencies and 2) the verb-noun co-occurrence patterns can discriminate noun meanings. This study investigated to what extent the verbs that co-occur with abstract nouns can predict their representation, as measured in pair-wise similarity ratings by a sample of human raters. Low-discriminating verbs should co-occur with all kinds and classes of abstract nouns with equal frequency. Such verbs are general and predicate the entire class of abstract objects. Verbs of medium distinctiveness are more useful for classification, since they group a subset of abstract nouns together while excluding other subsets. Finally, high-discriminating verbs co-occur with only a few abstract nouns. They provide the basis for finer sub-classification of abstract nouns.

Our word sample contained 30 nouns, selected to represent a range of ontological classes (e.g., Vendler, 1957) including process, state, intentional and unintentional events, emotions, and attributes. Our corpus consisted of 7500 sentences, with 250 sentences for each noun, randomly selected from LexisNexis. This database contains text from major newspapers and magazines, covering a wide range of topics. We established a matrix for every noun, plotting the co-occurrence of nouns and verbs. The cells of the matrix contain the co-occurrence frequencies. The resulting list of more than 1500 verbs was reduced by two independent raters to 37 verb classes based on the

classification provided by Levin (1993). For simplicity, we restricted the analysis on verbs that can be classified in Levin's system. Out of the classifiable verbs, the most frequently occurring verb classes were possession, creation / transformation, verbs with predicative complements, and verbs of communication, which accounted for 54% of the classified verbs. For each noun, the context analysis yielded a vector, with each element indicating the co-occurrence of verbs from each particular verb class and the noun. Each noun context vector contained the verb classes twice, once for the abstract noun in agent position, and once for the noun in object position, resulting in 74 elements.

We computed the cosines for all the 435 vector pairs to measure the verb context similarity of abstract nouns. Cosines vary from 0 to 1, with higher cosines indicating higher vector similarity. The cosines, measuring context similarity, were compared to similarity judgments of the corresponding abstract noun pairs. These similarity judgments were averaged across 33 raters (all college students). Verb vector cosines should be significantly correlated with the human ratings if the verbs used in the context of abstract nouns provide distinctive information about their meanings. To estimate the predictive power of the co-occurring verb in comparison to the entire context, we also correlated both measures with context vector cosines based on Latent Semantic Analysis (Landauer & Dumais, 1997), which represents a word's entire linguistic context. We used LSA vectors from two different text corpora, the encyclopedia (E) and general reading up to 1st year college (GR). The correlation of human ratings with LSA cosines may be higher if context information beyond verbs provides additional relevant information, or lower, if the additional information produces " noise " in the vectors.

Cosines of the verb context vectors correlated very little with the LSA cosines (GR: $r = 0.10$, $p < 0.05$; E: $r = 0.02$, $p > 0.05$). However, both the verb context vectors ($r = 0.22$, $p < 0.01$) and the LSA vectors (GR; $r = 0.23$, $p < 0.01$; E: $r = 0.17$, $p < 0.01$) were significantly correlated with the human ratings. Thus, verbs used in contexts of abstract nouns, as well as the overall context, are significantly related to the semantic relations among the nouns. It is interesting that the additional context information in LSA did not improve the correlation much. This result suggests that verbs carry a lot of the relevant information. This makes sense, since abstract concepts are often related to causal or temporal relationships inherent in situations and complex interrelations of agents. It is conceivable that verbs express such relations and therefore reflect the meanings of the abstract concepts accurately.

Can abstract concept classes can be distinguished based on context? We constructed a variable that indicated for each noun pair whether it belongs into the same ontological class (1) or not (0). A logistic regression tested to what extent this variable was predicted by the human ratings, verb vector cosines, and LSA cosines. The only significant predictor in the analysis was the human similarity ratings ($p < 0.01$). Thus, context similarity is related to semantic aspects of abstract concepts, but fails to discriminate between their ontological classes. Future work will analyze further context elements and examine the relative contribution of particular verb classes.

AutoCoder: An Intelligent Assistant for Coding Protocol Data

J. Earwood, M. A. Durbin & R. M. Golden

University of Texas at Dallas (U.S.A.)

JWEarwood1@aol.com

In many text comprehension experiments, human subjects are asked to either: (i) recall texts, (ii) summarize texts, (iii) answer questions about texts, or (iv) "talk-aloud" during their processing of a text. The resulting protocol data must then be "coded" using a coding procedure relying on

human coders. Although inter-coder reliability measures are often reported, such coding procedures are fundamentally problematic since they are difficult to quantitatively replicate and the explicit "coding theory" is not directly accessible to other researchers. This preliminary research aims to address these issues using a special user-friendly artificially intelligent graphical user interface called "AUTO-CODER" which is a software-instantiated psychological theory of how text clauses are mapped into complex propositions. Specifically, 24 human subjects read and verbally recalled the "miser text" which is a text consisting of approximately 14 complex propositions. Here is the "miser text":

A miser bought a lump of gold using all of his money. The miser buried the gold in the ground. The miser looked at the buried gold each day. One of the miser's servants discovered the buried gold. The servant stole the gold. The miser, on his next visit, found the hole empty. The miser was very upset. The miser pulled his hair. A neighbor told the miser not to be upset. The neighbor said, "Go and take a stone, and bury it in the hole." The neighbor said, "And imagine that the gold is still lying there." The neighbor said, "The stone will be as useful to you as the gold." The neighbor said, "When you had the gold, you never used it."

The human subject recall data was then transcribed and text segments corresponding to complex propositions were identified by a skilled human coder. Next, the skilled human coder mapped each text segment in the human recall protocol data from 12 of the 24 human subjects (i.e., "the training data set") into a complex proposition while AUTO-CODER simply "watched". More precisely, the human coder mapped selected key words in the text into selected key concepts and then assigned a complex proposition to a given concept subsequence and AUTO-CODER recorded: (1) the probability that one concept followed another in the "coded protocol data" assuming a particular complex proposition (i.e., the "transition probability"), and (2) the probability that a given concept would be expressed as a particular word or word subsequence (i.e., the "emission probability"). Next, the transcribed and coded recall data of the remaining 12 human subjects which constituted the "test data set" was then presented directly to AUTO-CODER without any additional coding assistance from the human coder. Using the classic Viterbi Algorithm in conjunction with a hidden Markov model methodology, AUTO-CODER used the statistics collected from observing the behavior of the human coder on the "training data set" to simultaneously infer the "most probable" assignment of complex propositions to text segments. In this preliminary version of AUTO-CODER, AUTO-CODER always generates the "most probable" complex proposition for a given text clause despite that complex proposition's improbability. Thus, "false alarms" (i.e., detecting a complex proposition to be present when in fact it is absent) are likely.

AUTO-CODER's performance was impressive. The percentage agreement that AUTO-CODER and a skilled human coder agreed that one of the 14 possible complex propositions was mentioned (or not mentioned) in a particular subject's text recall taken from the "training data" was 95% (Cohen's Kappa = 91%). For the "test data", the agreement measure was 83% (Cohen's Kappa = 65%). Although these measures of agreement seem slightly low, it is important to remember that AUTO-CODER is a computer program and so codes protocol data consistently in exactly the same way every time (i.e., AUTO-CODER has "perfect reliability"). The particular dictionaries stored in AUTO-CODER represent a theory of how the human coder codes the data. Thus, the 83% agreement measure and 65% kappa measure on the test data must be interpreted not as measures of "reliability" but really as measures of the "validity" of AUTO-CODER's behavior. AUTO-CODER is able to "solve" the natural language understanding problem this effectively because it's dictionaries are all text-specific, the text is very short, and AUTO-CODER is only searching for 14 complex propositions. By state-of-the-art standards in the field of artificial intelligence, AUTO-CODER works with "toy problems". However, it is important to emphasize that the problem of automatically and reliably coding human subject protocol data (although a "toy problem" for artificial intelligence researchers) is an extremely serious and important one for researchers in the field of text comprehension.

To obtain some idea regarding the richness of the human protocol data, here are example recall protocol data excerpts from the "training data set" and "test data sets". Concepts and propositions

are presented in UPPER-CASE LETTERS, while the human protocol data is presented in *lower-case italic letters*.

Text Segment ("training data") Coding Specified by Human Coder In order to "Train" AUTOCODER

someone that a servant that knew that discovered the money

Concepts: SERVANT, FOUND, GOLD Proposition: (FOUND, SERVANT, GOLD)

and took it

Concepts: STOLE, GOLD PROPOSITION: (STOLE, SERVANT, GOLD)

and then the miser saw that the money was gone

Concepts: MISER, MBUILD, GOLD, GONE Proposition: (MBUILD, MISER, "GONE(GOLD)")

and he was upset

Concepts: MISER, UPSET Proposition: (UPSET, MISER)

Text Segment ("novel test data") Coding Generated by AUTOCODER

he buried the lump of gold in a hole in his backyard

Concepts: SOMEONE, BURY, GOLD, HOLE Proposition: (BURY, MISER, GOLD)

and every day he would go and look at the money

Concepts: SOMEONE, ATTEND, GOLD PROPOSITION: (LOOK, MISER, GOLD)

in the hole

Concepts: HOLE Proposition: (TOLD,NEIGHBOR, MISER ,"ATTEND(STONE)") Poor AUTOCODER Inference!!!

and one day he went to look at the money in the hole

Concepts: SOMEONE, WENT, ATTEND, GOLD, HOLE

Proposition: (COMPLAIN, MISER, NEIGHBOR) Poor AUTOCODER Inference!!!

Conceptual Change via Learning from a Text. A Connectionist Model

V. Sanjose*, O. M. Padilla** & E. Vidal-Abarca*

*University of Valencia, (Spain)

**Automa University of Zacatecas (Mexico)

Eduardo.Vidal-abarca@uv.es

The instructional study. This study describes an experiment about conceptual change on scientific concepts via learning from a text and a connectionist model simulating that change. Thirteen students from the Teachers College at the university of Valencia participated in the study, though only data from three of them are analyzed here. Five instruments were used to get the students' prior background knowledge about some scientific concepts (i.e., heat, temperature, thermal equilibrium and conductivity): (a) an open-ended questionnaire, (b) a clinical interview, (c) a multiple-choice test, (d) a laboratory experience, and (e) the examination of an individualized conceptual-map made up of the participants' misconceptions.

Each students then participated in an individualized study session. First they read a text about the scientific concepts on a computer screen using the program Select-the-Text (Goldman & Saul, 1990). Most of the information in the text was common to all participants, but part of it was specific to each one, and it was aimed at contradicting specific personal misconceptions. Participants read the text at their own pace. The order in which paragraphs were read, and the time spent on each one were recorded. Then, students answered an open-ended questionnaire. Some questions were aimed at making apparent the contradiction between the specific personal misconception and the scientific explanations. Participants could re-read the text to answer the questions. In order to measure conceptual changes in the students understanding, they were administered a free recall test, and the same multiple-choice test they had done in the pretest phase.

Two out of three students revised their misconceptions considerably. They paid a lot of attention to the text information which contradicted their misconceptions. However some erroneous ideas persisted after studying the text which may be attributed to: (a) the complexity of some concepts (e.g., energy), (b) lack of comprehension and learning strategies (e.g., focus on unimportant information), and (c) failure in the teaching procedure (e.g., failure to include text information and questions to contradict specific misconceptions). The third student scarcely changed his erroneous ideas. He had strong misconceptions about the scientific concepts probably due to his job as a fireman. Moreover, he refused to change his ideas about the topic because "he had a lot of knowledge about the text topic" (sic). When reading the text he avoided reading information which contradicted his misconceptions, and he did not include any conflictive information in his recall protocol.

The connectionist model. Our model assumed that the processing system consists of a thickly interconnected net of units which take on values of activation depending on the weights associated to the connections between pairs of units (inner weights), and between those units and the text input (outer weights). The model deals with semantic units, i.e., propositions, of three types: Text propositions (P-units), propositions coming from the reader's prior background knowledge (K-units), and propositions that are a product of the reader's inferences (I-units).

When students read a text they form a net by building propositions from the text. Simultaneously, both the text and the propositions activate the reader's prior background knowledge. As a result readers add nodes from their prior background knowledge to the textual net and establish connections among them to build a coherent mental representation of the text. This process may produce inferences as well.

In our model the processing of the text occurs in cycles, each one corresponding to a specific sentence. Within each cycle, the textual input activates specialized units. After each reading, external weights are updated following the Kohonen algorithm making the units more specialized. These units activate either P-units, K-units, or both. This process promotes changes in the internal weights. Every unit receives exterior input (i.e., words in the text), and input from its neighbors (other propositions); then they compute the output (i.e., activation) according to the weights. This output is communicated to the units in the net to which it is linked

An activation rule was created that produced a level of activation ranging from 0 to 1. The nodes which are positively connected to many others in the net will be strengthened and their activation will increase. The nodes with few or negative connections will be weakened and their activation will decrease. The pattern of activation indicates the role that each node plays within the net, and may be considered as a measure of the importance of the meaning of each node in the mental representation at a certain time (Kintsch, 1998).

A learning rule was implemented to model how the weights in the connections are modified, i.e., how learning occurs. The adjustment of the weighs takes place in every iteration when activation changes, and it occurs simultaneously in all the connections in the cycle. Weights are adjusted so that internal connections reproduce the pattern of activation completed by the external input to that unit (McClelland, & Rumelhart, 1986); That is, the effect produced in the net by the text will be ideally reproduced without text. It would correspond to complete learning which does not take place in real subjects. The weights of connections among K-units are also modified as a consequence of the process. These changes are controlled by Hebb's rule. In the case of inhibitory connections, the anti-Hebb rule is used. Finally, for connections that produce inferences, a modification of Hebb's rule is used.

As in all connectionist models, only weights associated with connections are stored in LTM. Thus, when some signals of recovery are present, information storage allows the recreation of the patterns of activation originally produced by an external input. In our case, once the text is read, internal connections are stored in each cycle, so that at a later time, from a fraction of the input pattern, the subjects may recover virtually all of the activation pattern. For the recall test, we assume that readers take some of the concepts mentioned in the text as recovery signals and thus they reproduce the values of activation of a particular input, which must correspond, hypothetically, to what they write on the recall protocols.

When results of the simulation and data of recall protocols were compared, we found a coincidence of 79% for the two students who revised their misconceptions. Moreover, all the information actually recalled was included in the simulation. However, for the student who scarcely changed his misconceptions the coincidence was only 38.2%, although all ideas recalled by the student were included in the result of the simulation. This result illustrates the limitations of connectionist models. They can not simulate properly non rational behaviors. It seems that the last students approached the study of the text without a real desire to learn, which inhibited rational learning strategies. Connectionist models can not properly model these sort of non rational behaviors.

References

Goldman, S.R. & Saul, E.U. (1990). Applications for tracking reading behavior on the Macintosh. Behavior Research Methods, Instruments and Computers, 22, 526-532.

Kintsch, W. (1998). Comprehension: a paradigm for cognition. Cambridge, MA: Cambridge University Press.

Rumelhart, D. E., McClelland, J. L., and the PDP Research Group (1986). Parallel Distributed Processing. Cambridge, MA: MIT Press.

The Effects of Information Presentation Style on Question Generation

S. D. Craig, B. Gholson, M. Ventura & The Tutoring Research Group

The University of Memphis (U.S.A.)

scraig@memphis.edu

For many years, researchers have known that question asking is intricately involve in the cognitive processes used in such tasks as text comprehension (e.g., Collins, Brown, & Larkin, 1980; Graesser, Person, & Huber, 1992; Hilton, 1990; Kintch, 1998), and learning complex materials (Collins, 1985; Graesser, Singer, & Trabasso, 1994; Palincsar & Brown, 1984; Schank, 1986). There are two problems that have been identified in the literature: Students do not tend to ask questions, and the few questions they do ask are often unsophisticated (Graesser & Person, 1994). We explored these problems by manipulating the method of information presentation in a vicarious learning format. Two Microsoft agents were used to model information presented as either a monologue or dialog. The design of the experiment was a 2 (acquisition condition: Monologue vs. dialogue) x 2 (tutorial content: Software vs. hardware) factorial.

Method

Participants. A total of 48 students, whose participation met a course requirement, were drawn from introductory psychology classes at The University of Memphis and randomly assigned to one of the four groups. The participants included only students who reported that they had average or less knowledge of computers.

Materials and procedure. (1) Acquisition: During acquisition a virtual tutor and a virtual tutee, along with a picture relevant to the current subtopic, were always on the monitor. Half the students were presented with a virtual tutoring session involving eight subtopics with hardware content and half eight subtopics with software content. Half the students in each content condition observed a monologue and half a dialog. In both of these conditions each subtopic was introduced by a brief information delivery presented by the virtual tutor. In the monologue condition the virtual tutee then asked one question, and the virtual tutor followed this with a monologue presentation of that subtopic's tutorial content. In the dialogue condition, after each information delivery the virtual tutee asked a series of questions, ranging from five to 14 across the eight subtopics (66 total). The exact words, phrases, and sentences used by the virtual tutor in response to the virtual tutee's question(s) were identical in the dialogue and monologue conditions in each content domain. At the end of acquisition, each student was given an open ended retention test on two of the subtopics, balanced so that each subtopic was presented to three different students. (2) Transfer : In transfer the students were told that the tutor would deliver some information on each subtopic, and then they (the student) could clarify any information that would help them understand the subtopic. Only the virtual tutor and the appropriate picture were on the monitor for each of the eight subtopics. Students in the software content condition during acquisition were presented with hardware transfer, and those in hardware content condition transferred to software. The contents of the information deliveries for each subtopic in transfer were identical to those

presented for the corresponding subtopic in acquisition. Transfer began with the virtual tutor giving the information delivery for the first subtopic. The virtual tutor ended the information delivery by inviting the student to clarify anything that would help them understand the material, but then the tutor stopped talking. The student then queried the experimenter until they said they were finished with that subtopic (i.e., said they had no more questions). The virtual tutor then brought up a new picture, gave the information delivery for that subtopic, and so on. After the eight subtopics were presented in transfer, each student received a second retention test that, parallel to acquisition, consisted of open ended questions about two of the subtopics covered in transfer.

Results and Discussion

In the retention tests students in the dialog condition wrote significantly more content (48.9 propositions) than those in the monologue condition (36.5 propositions), $F(1, 44) = 6.85$, $p < .05$. With respect to the total number of turns taken, again students in the dialogue condition outperformed those in the monologue condition (dialog = 31.4 turns, monologue = 19.0), $F(1, 44) = 6.38$, $p < .05$ (Craig, Gholson, Smither, & The T.R.G., 1999). We categorized queries by the participants using the Graesser, Person and Huber (1992) taxonomy for questions. This taxonomy includes a variety of different categories, which are organized into questions requiring short and long answers. A preliminary test of independence revealed that questions asked in the monologue and dialog condition were not equally distributed across question types, $X2(df = 17) = 165.55$, $p < .001$. We then performed a 2 (training condition: Monologue vs. dialogue) x 2 (tutoring content: Hardware vs. software) x 12 (question category) multivariate analysis of variance. This analysis revealed a significant interaction between training condition and question category, $F(1, 44) = 5.70$, $p < .001$. In follow-up analysis, significant differences between the monologue and dialogue conditions were found in eight question categories. Students in the monologue condition exhibited significantly more shallow level queries from verification, concept completion, and assertion categories (see Graesser & Person, 1994). Those in the dialogue condition generated more deep-level questions from comparison, example, interpretation, causal consequence, and enablement categories.

Students in the dialog condition learned more, took more turns in mixed-initiative dialog, and asked more questions than those in the monologue condition. The way information is presented to listeners in a vicarious-learning format (McKendree, Stenning, Mayes, Lee, & Cox, 1998) has an impact on the way they interact via question asking. If a question-rich dialogue is used to present information, students ask more and deeper-level questions. This deeper-level question asking led to better mental representations of the topics and, thus, increased learning.

References

Craig, S. D., Gholson, B., Smither, D. & The T.R.G. (1999). Listening in on dialogues and monologues of embodied agents in virtual tutoring sessions: Learning and questioning. Presented at The 9th Annual Meeting of the Society for Text and Discourse. Vancouver, Canada.

Collins, A. (1985). Teaching and reasoning skills. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), Thinking and learning skills, Vol. 2 (pp. 579-586). Hillsdale, NJ: Erlbaum.

Collins, A., Brown, J. S., & Larkin, K. M. (1980). Inference in text understanding. In R. J. Spiro, B. C. Bruce, & W. F. Brewer (Eds.), Theoretical issues in reading comprehension (pp. 385-407). Hillsdale, NJ: Erlbaum.

Graesser, A. C., & Person, N. K. (1994). Question asking during tutoring. American Educational Research Journal, 31, 104-137.

Graesser, A. C., Person, N. K., & Huber, J. D. (1992). Mechanisms that generate questions. In T. E. Lauer, E. Peacock, & Graesser (Eds.), (pp. 67-187). Hillsdale, NJ: Erlbaum.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. Psychological Review, 101, 37I-395.

Hilton, D. J. (1990). Conversational processes and cuasal explanation. Psychological Bulletin, 107, 65-81.

Kintsch, W. (1998). Comprehension: A paradigm for cognition. Cambridge: Cambridge University Press.

McKendree, J., Stenning, K., Mayes, T., Lee, J., & Cox, R. (1998). Why observing a dialogue may benefit learning. Journal of Computer Assisted Learning, 14, 110-119.

Palincsar, S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. Cognition and Instruction, 1, 117-175.

Schank, R. C. (1986). Explanation patterns: Understanding mechanically and creatively. Hillsdale, NJ: Erlbaum.

Advertising Discourse: A Critical Approach to Chinese Advertisements

V. T-H. Chang

Universidad do Minho (Portugal)

vthchang@ms32.hinet.net

This paper aims to propose a critical study of Chinese advertisements. Data for analysis are chosen on the grounds of the remarkable trends conveyed in these advertisements. First, most of them do not just focus on one or two slogans, as compared to earlier ones, but on much longer texts or storytelling, and this has been widely employed in many other commercials and political propaganda in recent years. Secondly, they adopt the linguistic devices of metaphors, puns, irony, and poetic effects. These particular figures of speech and stylistic patterns make the advertisements attractive and somehow literary, thus arousing the attention of the audience, and serving a site where the weaker effects of communication of feelings, attitudes and impressions could be more implicated than by literal forms of expressions in strong communication. Thirdly, they are ideologically significant. The bookstore no longer merely sells books, the department stores are no longer limited to selling clothes that are in vogue, and the political propaganda is beyond patently demonstrating their political claims. Instead, by reflecting modern lifestyles and current social concerns therein, they evoke popular frames such as individualism, intellectualism, elitism, communicative and political issues, humanistic concerns, feminine awakening, corporate imaging, indigenization, group interests and those experiential values of petits bourgeois. They add sociocultural properties to the commodities being advertised and lead the audience to recognize the prominent values.

The speaker-audience relationship in communication is reexamined from looking into thirteen Chinese advertisements. On the speaker's side, she could be exempt from endorsing the assumptions implied and from heavier social responsibility by exploiting these tropes in verbal communication, leaving much wider interpretation space accompanying larger responsibility to the audience. On the audience's side, he searches for optimal relevance in the interpretation process during which a wide array of weak implicatures, based on his greater share of responsibility, would be derived from the non-literal speech together with the context, depending on the different degrees of involvement and shared cognitive environment. Critical discourse analysis, on the other hand, rendering the sociocultural perspective of language use, is further applied to investigate the inseparable relationship between language and social meaning. Traditional account of Eastern versus Western cultures falls into the general dichotomies: Introvert versus extrovert, passive/submissive versus active/aggressive, self-denying versus self-assertive, involvement versus independence, power versus solidarity, hierarchy versus equality. But different cultures often put different degrees of emphasis on these values and their worldviews tend to be comparatively different in some respects. Those longstanding traditional Chinese values facing challenges and undergoing wide-ranging cultural impacts via the above-mentioned prevalent frames, with the social continuity and change originated from competing values and intercultural communication, have also been explored in this study. For instance, "MOTOROLA cd928 makes people indulge in sound and colours." Reading the word *indulge in sound and colours*, a pun used by the advertiser, the addressee will probably recover the meaning of " indulging in carnal pleasure and sensory enjoyment ", which would be the most accessible interpretation for its idiomatic usage. But it will have to be rejected as inconsistent with the principle of relevance (Sperber & Wilson 1995), cause that is a caption for cellular. The addressee will then recover the propositional form identical to its literal meaning, "MOTOROLA cd928 makes people indulge in a sphere of sound and colours. " Interesting is the case that, after considering the relationship between language use and social function, whether the abandoned interpretation belongs to nonsense, actually contributing to the intended interpretation, or a meaning also actually communicated by advertiser, as classified by Tanaka (1994), is not so definitely clear-cut. Carnal pleasure and sensory enjoyment are in general loaded heavily with socially negative impressions. "The more you indulge in such environment, the faster you will be ruined. " is one of the traditional and powerful value judgments under Chinese cultural norms. However, the sensory enjoyment and even carnal pleasure could have become one frequent negotiation activity in an informal business setting, representing sophisticated charisma and constituting another iconicity of corporate culture. It is not weird to see three or five businesspersons talking business in such places. Moreover, like many vague items, *indulgence* involves the problem of "degree" which, together with the pun, helps advertiser evades larger responsibility. Not merely the conventionalized and conversational implicature are competing, but the traditional social ethics and cultural identity are facing challenges. This vague and sensational wording used here are to trigger an imaginative space for the audience living in such "a cement jungle" preoccupied with annoying work and routine life. It compensates them with a tantalizing mood in accord with the lives of petits bourgeois, weakly communicating such implicature as, "If you're not able to go to these places and attend such activities, it's OK, just stay with MOTOROLA cd928, it'll also bring you a colourful lifestyle." In this sense, it seems that both meanings are actually communicated. Thus, from socially cognitive aspects, whether the implicatures inferred from "carnal pleasure and sensory enjoyment" could outweigh those from "sound and colours" provids another site where emergent two competing illocutionary forces.

Promotion of products, services and images is the ultimate goal of advertising. The selling motive, however, as illustrated in this study, has largely been hidden or melted by persuasion through storytelling and experience-sharing in an ever increasingly prominent trend. Aside from lexical and phrasal hedges, macro-level hedging via social metaphor is also initiated intertextually to shift consumers' buying concern and voting action. Such ambivalent language-use strategy not only echoes the traits of indirectness and implicitness of Chinese culture, thus maintaining the harmonic relationship of face-saving and mutual satisfaction, but conveys dominant ideologies invisibly and effectively due to its vagueness. Advertising, lending itself as a symbolic domain of dominant ideologies (Lull 1995), not merely incorporates fashion-driven discourse by reflecting social cognition and cultural patterns, but invites the audience members as potential consumers to

recognize these prominent values. Advertisements are socioculturally shaped but they also constitute sociocultural cognition, in ways that may be transformative as well as reproductive (Fairclough 1995), thus keeping the dialectical relationship between social structures and social practice/discourse. This functional and critical linguistic study of advertising texts heightens sensitivity to language and its (mis)uses, poses a preliminary stage to cross-cultural communication and interdisciplinary studies, which themselves constitute promising research issues to overcome the symplistic account of Eastern versus Western cultures, and provides useful insights into a multitude of issues, including symbolic power, social stereotypes, changes in social trends and attitudes, group and ethnic identity, cultural identity and iconicity, verbal art, and gender.

Language-Related Causes of Communication Breakdown in Medical Practice

M. Ibba

Universita Cattolica del S. Cuore, Roma (Italy)

m.iblangmed@rm.unicatt.it

Crucial to the validity of signed informed consent forms in hospital medical practices, such as diagnostic procedures, therapeutic courses, surgical treatment, research protocols and the like, is the understanding of the information that is due to be disclosed to the patients and/or to research subjects through both the printed materials – informed consent forms and other informative literature – and through the health care provider-patient / research subject verbal exchanges leading to the actual consent to treatment and/or experimentation. Lack of understanding of the information thus provided is reported to have often been the cause of litigious attitudes on the part of those patients, or of their guardians, who felt that their information rights were violated, with consequent law suits for malpractice for negligence in disclosure of crucial information.

Close analysis of a number, and of a variety, of consent forms and of other printed information literature, coupled with the observation and recording of actual dialogues between physicians and patients, between researchers and research subject, in two major university hospitals in the Unites States of America, have led to the identification of most commonly recurrent language/discourse features which proved to be predominantly responsible for the patients inadequate awareness of, for instance, the risk/benefit ratios entailed by the proposed procedures and treatments, together with other crucial information.

Medical jargon, deemed appropriate in verbal exchanges among health care providers, was found to be one of the obstacles to proper understanding of the messages addressed to the often layman patient. The use of abbreviations and of achronims was also a frequent source of patients' disoriented, dissatisfied and, consequently, not compliant behaviour. The frequent use of metaphors and of euphemisms in medical speech also interfered with clear reception of the messages. These and other language-related communication problems were observed. The findings of the study urged various, successful attempts on the part of this researcher to seriously engage in professional dialogues with the physicians, with the nurse, with the research directors, with the protocol review team members, in order to assist them in some reflection upon the undeniable pitfalls in the language, written and spoken, used by them to convey important information to the recipients of their professional care.

As for the written information materials, specifically, copies of edited versions of the existing consent forms, brochures, and leaflets, were provided to the personnel in charge and formal requests were issued for the recognized, much needed revisions of such materials that were routinely provided to patients. Besides the strictly linguistic problems hindering communication,

sociolinguistic factors as well proved to constitute additional impairments to successful verbal transactions. Furthermore, the psychological state of hospitalized patients and the research subjects was inevitably observed, and interesting annotations were made. Yet, an analysis of this aspect of the study was deliberately left out, because it was understandably considered territory of inquiry for scholars with specific expertise and research tools.

All the observations performed by this research were always preceded by request of and freely given consent by the people involved: Health care providers, patients and hospital administrators. Needless to say, many questions remained unanswered, in the study reported in this paper. They remained as research material urging further inquiry.

More than a Stately Dance: Dialogue as a Reaction Time Experiment

E.G. Bard, M.P. Aylett & M. Bull

University of Edinburgh (Scotland)

ellen@ling.ed.ac.uk

Most discussions of the timing of turn-taking in dialogue concentrate on the new speaker's sensitivity to prosodic or syntactic characteristics of the last speaker's utterance which predict where that speaker intends to finish (Ford & Thompson, 1996; Sacks et al., 1974; Wells & Macfarlane, 1998). Dialogue is treated as if it were a pas de deux, with timing of one partner's activity aligned to timing of the other's (Couper-Kuhlen, 1993), and with interpersonal considerations important to each partner's response to the other's timing signals (Warner, 1992). In monologue, however, the timing of speech onsets is often treated as evidence for models of language production (Beattie, 1983 ; Butterworth, 1980 ; Wheeldon & Lahiri, 1997), with pauses determined by the complexity of the unit being planned. In other experimental paradigms, time to begin speaking is used to study perception, comprehension, or problem solving (Bock, 1996). If time to begin speaking in dialogue is also subject to these cognitive effects, then " switching pauses " should behave like reaction times. We make use of a large corpus of task oriented dialogues to examine delay, both positive and negative, from the offset of one speaker's utterance to the onset of the next's. We show that even with basic social considerations taken into account, these intervals are sensitive to the cognitive and production tasks currently presented to the new speaker.

Our observations are based on the HCRC Map Task Corpus (Anderson et al., 1991), 128 dialogues from 64 pairs of undergraduates engaged in a route communication task. Instruction Giver's and Follower's maps matched only in alternate landmarks. Participants knew that their maps differed but not where or how. Each participant served as Instruction Giver for the same route to two different Followers and as Instruction Follower for two different routes. The corpus was balanced for familiarity of participants and for ability to see the interlocutor's face. Digital stereo recordings were segmented at word boundaries, and coded for landmarks mentioned, for Dialogue Moves and Games (sub-tasks of the dialogue itself, Carletta et al., 1997), as well as for Transactions (sub-sections of route communication task which are executed by sequences of Games (Isard & Carletta, 1995). Deviation of drawn from model routes in sq-cm reflected task performance. Since " turns " in the usual sense were coded as one or more Dialogue Moves from a single speaker, time to begin speaking is found in the intervals between successively starting Moves by different speakers.

From the 18.248 such intervals in the corpus, those not likely to separate a Move and its response were excluded, using temporal limitations established by pretest: The second Move started too soon (less than 350 msec) after the first or overlapped the first too much (began more than 1 sec

before the first Move's offset). BASIC data come from half the corpus (5546 intervals) where video recordings allow elimination of intervals in which drawing occurred. Multiple regression analysis tested main effects and interactions for social, cognitive and production variables derived from design and coding. The resulting equations were applied to the REMAINING dialogues (7444 intervals) and then to both halves together. We report results only where Beta-values achieved $p < .05$ in the basic materials and overall and $p < .10$ in the non-basic while keeping the same sign throughout. The proportion of significant contributors was well above chance.

Interpersonal effects. If between-speaker Inter-Move Intervals (IMIs) are sensitive to relationships between the speakers, they should show effects of channel (faces visible or hidden), familiarity, or sex differences. In fact, Different-sex pairs produce shorter IMIs than same-sex pairs if the interlocutors have just met, but not if they are already friends.

Cognitive Effects. If IMIs are reaction time measures, they should show the kinds of generic effects usually found for RTs in problem solving tasks. In fact, IMIs should show independent contributions from 3 factors which normally affect problem solving. There is a practice effect, with shorter IMIs later in the dialogues; a task difficulty effect, so that dialogues in highest quartile of route error scores have longest IMIs; and a speed x accuracy trade-off, with IMIs longest in the dialogues with lowest error scores. If IMIs are RTs in on-line language comprehension tasks, they should reflect the difficulty of comprehending the preceding Move. IMIs are significantly shorter after reference to discourse-Given landmarks whether or not they appear on the responder's map. IMIs are longer after references to New landmarks which the responder lacks but just in the case that they are introduced with no Given shared landmark. If IMIs also reflect the utterance production process, as pauses in monologue are said to do, they should be sensitive to the complexity of the utterance which is about to be spoken. In fact, IMIs are longer before longer utterances. IMIs also increase in length with the scope of the higher level dialogue unit which the utterance initiates: They are shortest before Game internal Moves, longer before Game-initial moves, and longest before Transaction-initial moves. All these simultaneous effects are independent. Their relative contribution to the explained variance in IMIs promises a new way to categorize spontaneous spoken language genres: in terms of the relative importance of social, task, comprehension, and production considerations to the speaker's RT.

References

Anderson, A. Bader, M., Bard, E.G., Boyle, E., Doherty, G., Garrod, S., Isard, S., Kowtko, J., McAllister, J., Miller, J., Sotillo, C., Thompson, H. S., & Weinert, R. (1991). The H.C.R.C. Map Task Corpus. Language and Speech, 34, 351-366.

Beattie, G. W. (1983). Talk: An anlysis of speech and non-verbal behaviour in conversation. Milton-Keynes: Open University Press.

Bock, K. (1996). Language production: Methods and methodologies. Psychonomic Bulletin & Review, 3, 395-421.

Butterworth, B. (1980). Evidence from pauses in speech. In B. Butterworth (ed.) Language Production, vol 1. London: Academic Press.

Carletta, J., Isard, A., Isard, S., Kowtko, J., Newlands, A., Doherty-Sneddon, G., & Anderson, A.H. (1997). The reliability of a dialogue structure coding scheme. Computational Linguistics, 23, 13-31.

Couper-Kuhlen, E. (1993). English speech rhythm: form and function in everyday verbal interactions. Amsterdam: John Benjamins.

Ford, C.E., & Thompson, S. A. (1996). Interactional units in conversation: syntactic, intonational, an pragmatic resources for the management of turns. In E. Ochs, E. Schegloff, and S. A. Thompson (eds.)Interaction aand Grammar. Cambridge: CUP.

Isard, A., & Carletta, J. (1995). Transaction and action coding in the Map Task Corpus. Research paper HCRC/RP-65. HCRC, University of Edinburgh.

Sacks, H., Schegloff, E., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. Language, 50, 696-735.

Warner, R. M. (1992) Dimensions of social-interaction tempo - a factor analytic study of time and frequency-domain indexes of interaction structure. Psycholinguistic Research, 21, 173-191.

Wells, B., & Macfarlane, S. (1998). Prosody as an interactional resource: Turn-projection and overlap. Language and Speech, 41, 265-294.

Wheeldon, L., & Lahiri, A. (1997). Prosodic units in speech production. JML, 37, 356-81.

Literacy is More than Learning to Read: Information Retention from TV News

S. Gulgoz, F. Goksen & C. Kagitcibasi

Koc University (Turkey)

sgulgoz@ku.edu.tr

The aim of the current study was to investigate the indirect discourse processing gains from literacy training. We argue that learning to read initiates a chain reaction of changes in the cognitive processing system. One type of change has been observed in the laterality of the brain. Lecours, Mehler, Parente and others (1988) have shown in their comparison of literates with illiterates that the left hemisphere dominance of language representation was more pronounced among schooled subjects. However, the changes are not restricted to functional specializations in the brain because reading and writing are more than individual skills; they are skills that encompass a number of more fundamental skills. Acquisition of such skills and knowledge, ranging from orthographic to syntactic, semantic to representational, has implications for other cognitive processes. For instance, memory processes are provided with a new tool such as writing, and a new medium of representation such as visual images of words that could be used in the service of remembering. While societies prior to literacy has been dependent upon oral traditions for both personal and public memories (Rubin, 1995), writing alleviated the pressure on cognitive resources of individuals. Nevertheless, Olson (1996) claims the effect is much more significant than that. As people become readers, they lose their ability to represent language in any other form than the form provided by the model of written language. Learning to read and write is a series of discoveries; discovery of the phonemic units, discovery of words, discovery of grammatical form, and discovery of sentence meaning (Olson, 1996). Therefore, with the introduction of written language, the representation of language becomes a representation that is largely different from the representation generated by oral language.

News is a special case of discourse. It is text read aloud, meaning that it is different from conversation but is also more restricted than reading. Processing strategies are more critical in listening to text read aloud because unlike reading, there is no chance to go back and re-read,

especially if the coherence of the text is low. Many illiterate people may lack the skill of reading but they may have more developed strategies for comprehending text read to them. They have to rely on TV news more than literate people do. In addition, they experience text being read to them more often because in daily life, when they need the information in a text, they tend to get someone to read to them. In order to investigate the rate of recall from text that was heard among illiterates and observe the effects of literacy training, we investigated the recall of information from TV news presented before and after attendance to a literacy-training course in the current study.

Method

The participants were 65 women with ages ranging between 17 and 65 with no ability to read or write at the beginning of the literacy training courses. The literacy-training program is a service of the Mother Child Education Foundation and is a 120-hour program conducted over three months, meeting three times a week. The emphasis is on the training of functional literacy taking advantage of the linguistic characteristics of the Turkish language (Durgunoglu, Oney & Kuscul, 1996). The news used in the study was created by making modifications on times, places, locations, and the names of real news and was recorded professionally in a national TV studio. The news were arranged to include different scopes (local, national, and international) and different types of news (hard news and soft news). When the scope and type of news were orthogonally varied, there were six types of news with one example of each. The news stories were recorded twice. In each recording of the news one example was recorded to contain the version of the news revised for coherence and the other example to contain the original version. The same person read the news in both versions without any visual effects. Women were presented the news program first and then interviewers approached them and asked them to recall first the topics of the news then the contents. They viewed the program twice with an interval of three months between them. Before the analyses, we wanted to check for any carryover effect from the previous viewing of the news stories. Prompting for recall before the second viewing did not reveal any recall. We also compared the participants who were present and absent during the first presentation of the program. This comparison showed no differences (F < 1). With the amount of interference from real TV news, this outcome was hardly surprising.

Results

The analyses indicated that women remembered more information after the training than before the training. This finding of improved memory was consistent in both the news topics and the news contents. The version of the news text also had an effect on recall of news topics such that the version revised for higher coherence was remembered better than the original version. This result was qualified by the interaction of version with the time of testing. Before the literacy training, there was no recall difference between versions but after training there was a significant difference favoring the revised version. The results of the content recall replicated this finding. The participants recalled the revised version better overall. However, there was a significant interaction between time of testing and text version. The interaction showed that the revised version was recalled better particularly after the training. Recall of the revised version after the training was significantly above the recall before literacy training. The analyses including news type and news scope revealed that recall increased after the training in all types of news with all the scopes used. There were interactions showing that the effect of training was most pronounced in local soft news and national and international hard news. The recall of content also showed a consistent increase in recall of all news types as well as the effects of news scope and news type. The interaction between scope and type showed the content of soft news to be recalled better in local and international levels while in national news the hard and soft news are recalled at equal levels.

Discussion

We concluded from our findings that literacy training provides a different way of processing information that prevails over spoken texts as well. Comprehension of spoken discourse is altered considerably by the introduction of written language system. The susceptiveness of the participants to text version indicates the importance of coherence especially in spoken text. The increase in news comprehension and the particular interactions related to the scope and the type of the news cannot be interpreted without reference to the activities involved in literacy training. The courses that were the sources of our samples employed a specific literacy-training program that emphasized functional literacy. During the course, women were taught to read material that came from real life, including price tags, instructions, bus schedules, and most significantly, newspaper headlines. This, no doubt, sensitized the participants of the literacy-training program to news items as well as improving their knowledge base for issues and people that take place in the news.

Functions of Honorifics and Topic Manipulation in "Discourse Politeness" as Unmarked Politeness: From the Analyses of Japanese Conversations

M. Usami

Tokyo University of Foreign Studies (Japan)

Usami@aol.com

This study analyzed 72 Japanese dyadic conversations between new acquaintances with regard to the choice of honorifics at the sentence level, and the shifts between the use and the non-use of polite forms within the same conversation and topic management strategies as discourse-level phenomena. I define "Discourse Politeness" as "a dynamic whole of functions of any element in both linguistic forms and discourse-level phenomena such as speech-level shifts and topic management strategies that play a part in the pragmatic politeness of a discourse." Focusing on these sentence - and discourse - level phenomena, I investigated situational variations in their use among dyads of speakers who vary in age and gender, important sociolinguistic variables in Japanese culture. More specifically, I examined the relationships between the sentence-level choice of honorifics, and the speech-level shift and topic management strategy as discourse-level phenomena. By doing so, I analyzed the dynamic functions of these elements produced in the overall Discourse Politeness in this group of conversations.

The results revealed six points about the present use of Japanese honorifics and the manipulation of speech level shifts and topic initiations. (1) The use of honorifics functioned more as a stylistic choice for the speaker than as a reflection of the actual relationship with the interlocutor as traditionally explained. (2) Female interlocutors used significantly more honorifics than did males; this was the only gender-related result. (3) Only the use of the non-polite form, a minor deviation from the normative language use and dominant speech level in the setting of this study, clearly reflected the age/power relationships between speaker and hearer. (4) The percentages of types "downshifts," that is, "downshift from self," "downshift from interlocutor" and "total downshifts", and of "upshift from interlocutor" as discourse-level phenomena clearly reflected the age/power relationships between speaker and hearer. In other words, speakers tended to downshift more frequently when talking with younger interlocutors, and to upshift from their interlocutor more frequently when they were talking with older interlocutors. These findings were statistically significant, as shown by an ANOVA. (5) Non-marked utterances that did not include any linguistic politeness markers accounted for about 25-30% of the total number of utterances in each conversation. This finding may indicate that speakers used non-marked utterances to avoid

acknowledging the speakers' asymmetric relationships by choosing linguistic politeness markers that already embedded the vertical human relationships. Thus, they may have functioned as a discourse politeness strategy. (6) The proportion of topic initiations between the speakers conspicuously reflected the power relationships between speaker and hearer. There were no significant differences in the manipulation of topic depending on either the interlocutors' or the speakers' gender.

Based on these findings, I return to the universal theory of politeness proposed by Brown and Levinson (1987), and claim that it is necessary to incorporate the concept of "Discourse Politeness" into studies of politeness. This will enable researchers to contrast politeness behavior in languages with and without honorifics within the same framework, and to construct a more comprehensive universal theory of politeness.

References

Brown, P., & Levinson, S. (1987). Politeness: Some universals in language usage. Cambridge University Press.

The Chechnya War: Soviet Political Discourse Revisited

T. Tsyrendorjieva

Moscow State Pedagogical University (Russia)

tattsyr@hotmail.com

Soviet political discourse (SPD) is identified as a part of social discourse determined by specific social and political conditions. This kind of discourse is characterized by a specific use of language mechanisms that is strategically aimed at ambiguity and ambivalence of meaning (Seriot, 1985) and triumphant demagogy (Zemskaya, 1996). Nominalized groups rather than propositions are used throughout the Communist party documents, reports, newspaper articles or TV news.

"The main source of labor production must be technological increase on the basis of developing and implementing new machinery and technological processes, intensive use of complex mechanization as well as deepening of specialization and improving cooperation". Semantically these countless nominalizations (gerund or a noun instead of active verb) lead to elimination of a subject, an agent. All the processes and actions are done without any active agent that helps manipulate the content making it seemingly meaningful. The demagogic language use favors a great number of cliches directed to some abstract or non-existing referent, euphemisms, non-direct nominations, and violations of all maxims of cooperation for the sake of linguistic manipulations. As a result, semantic elimination of the subject directs statements from an "ideal speaker" to an "ideal addressee".

The fundamental ambiguity and ambivalence of SPD in the 70s and demagogic language use are again strategically employed in the Russian political discourse of the 90s. The official and non-official Russian press to cover the Chechnya war uses the manipulative strategy. The collective and abstract referents such as 'band groups', 'terrorists', 'non-humans', 'threat to the community and

the society', and the representation of the one side as inherently positively marked, such as 'Russia as the main actor of the battle with the international terrorism', 'Russian troops as carrying freedom and order' substitute actual actors and actions by the familiar mythic enemies and defenders of a just society.

"Russia is fighting against the aggression of international terrorist organizations and Chechen bands that are struggling for the foundation of a new Islamic state in the south of Russia".

The main function of the Soviet ideological language as the function of ideological prescriptions was actualized discursively. The system of ideologemes was transformed into the "above text" of mythologemes as the discursive production of ideology was constructed mainly through the so-called Soviet ideological mythology. Myth is, according to Ivanov, a " form of speculation that goes beyond the scopes of speculation by that it wants to engender (give birth to) the truth it proclaims" (Ivanov, 1984, p. 29). The basic components of the Soviet ideological mythology were based on the concepts of collectivity and unity, progressivity, equality, freedom and justice. The most common of "hollow" ideological mythologemes was a mythologeme of the path (successful start of; accelerated, shock finish; finishing line) combined with a mythologeme of struggle (to see a target; to distribute forces; to throw all the forces for the struggle with (for the breakthrough); uncompromising struggle with) and a mythologeme of success (successful realization (decision, implementation); great creative stimulus, broad horizons of creation).

"Regressive forces are acting…In Chechnya we face them as well as the attempt to throw the country back to feudalism, slavery. It is an attempt to make history go back…" (God planety, 1996, p. 45). The SPD mythologeme of path is actualized as the threat of stepping back.

The analysis of Russian print press shows that democratic principles of humanistic discourse of the post-soviet period are competing with the demagogic properties of the Russian political discourse but are overshadowed by the discourse of othering/'enemying' and threat. 'Underground' parody and irony that opposed to SPD during the Soviet period is replaced by the discourse of patriotism and revitalized Russian national idea.

The analysis of political television program of interviewing shows the mixed character of discourse practice that bears characteristics of a conventional interview, performance, simulated conversation (Fairclough, 1995). The "mediatized" nature of distribution and consumption of such discourse practice makes intertextual and interdiscursive clash subtler and competing discourses less transparent although the constituent features of discourse practice speak to the revitalizing of the soviet discursive mechanism of manipulation.

References

Fairclough, N. (1995). Critical Discourse Analysis: the Critical Study of Language. Longman Group Limited.

Sériot, P. (1985). Analyse du discours politique soviétique. (Cultures et Sociétés de l'Est 2). Paris: Institut d'Etudes Slaves.

Zemskaya, E.A. (1996). Klishe 'novoyaza' I tsitatsiya v yazike postsovetskogo obtshestva. Voprosy yazikoznania, 3, 23-31. (Cliches of 'newspeak' and quoting in the post-soviet society language).

Understanding the Ironic Language

C. Munch

University of Paul Valery (France)

munch@alor.univ-montp3.fr

In the context of the understanding of ironic language, there are three different hypotheses with regard to the moment of intervention of the co-text (linguistic context). Giora (1995) upholds exhaustive access models which postulate a phase of exhaustive activation of all acceptations of a word or phrase followed by the later intervention of a co-text allowing its proper meaning to be grasped. Gibbs (1994) upholds selective lexical access models, according to which the subject has direct access to the ironic meaning of a statement without processing its literal sense. Lastly, Clark and Gerrig (1984) and Giora (1997, 1998, 1999) rely on Grice's conversational norms to postulate that access to the ironic meaning depends on the frequency or the salience of the acceptations of the words in the language.

Three hundred and eighty-four students took part in two experiments using tasks of semantic judgment and lexical decision. The results point to a process of exhaustive access to the various meanings of an ironic phrase followed by a selection of the proper meaning brought about by the co-text.

Experiment 1

Material. The experiment was carried out with pre-tested texts and targets. The subjects were presented with texts of three sentences: 1) introductory sentence; 2) a co-text sentence inferring an interpretation which is either literal (literal co-text) or ironic (figurative co-text) or both (ambiguous co-text); 3) a sentence to interpret. Each text was followed by a target which was literal (implying a literal interpretation of the text), figurative (implying an ironic interpretation of the text), non-associated (a French word not semantically associated with the text) or a nonword (group of syllables not forming a French word).

Procedure.

1) The subject selection stage: Each subject was presented with seven texts, four of them ironic, in a word-by-word Auto-Presentation segmentée (APS) procedure with mobile windows. At the end of each text, the subject saw the target between three stars and had to say if the target did actually qualify the last word in the text by pressing a yes-no key. Only subjects who replied correctly to at least three ironic texts out of four continued with the experiment.

2) The practice stage: The subjects were presented with four word-by-word Rapid Serial Visual Presentation texts with mobile windows. They were each given a single RSVP mode: 200, 300, 700 or 1000 ms per word. The target appeared between stars after each text with an Inter-Stimulus Interval (ISI) equal to 0 and during the same time than the other words. Subjects had to respond as rapidly as possible by pressing a key to say whether the target did actually qualify the last word in the text.

3) The experiment stage: Each subject was given twelve texts with an RSVP condition. They each saw four in literal co-text, four in figurative co-text and four in ambiguous co-text. Of the twelve texts, they saw three with a literal target, three with a figurative target, three with a non-associated target and three with a nonword target. Three texts were presented with an ISI = 0 ms, three with ISI = 100 ms, three with ISI = 300 ms and three with ISI = 1000 ms. The subject had

to say as rapidly as possible whether the target between stars at the end did actually qualify the last word in the text.

Subjects. One hundred and ninety two native French speakers between the ages of eighteen and thirty underwent the entire experiment.

Hypotheses. H1) In an exhaustive access model, all acceptations of the ironic phrase (i.e. both its literal and ironic meaning) are activated in the very first stages of the process (up to 300-350 ms according to Kintsch & Mross, 1985). It is only in a second stage of the process that the subject selects the proper meaning from the co-text (between 350 and 500 ms). Thus for 200 and 300 ms per word RSVP, in a figurative co-text, the literal and figurative targets will be activated with equivalent response times. For 700 and 1000 ms RSVP, the figurative target is accepted more widely and rapidly than the literal one.

H2) In a selective access model, only the acceptation which is consistent with the preceding co-text is activated and is so from the very first stages of the process.

For a figurative co-text, the figurative target is accepted more widely and rapidly than the literal one however long it is presented.

H3) In a selective access model based on frequency of acceptation in the language, the acceptation activated in the very first stages of the process is the one occurring most frequently in the language. Only if it is inconsistent with the co-text will the subject set about seeking other acceptations. For 200 and 300 ms RSVP, in a figurative co-text, the literal target is more widely and rapidly accepted than the figurative one. For 700 and 1000 ms RSVP, the figurative target is accepted more widely and rapidly than the literal one.

Results

The results point to a process of exhaustive access to the various meanings of an ironic phrase followed by a selection of the proper meaning brought about by the co-text. This last stage of selection is longer than the one observed in literal language.

Experiment 2

Material, procedure and hypotheses were the same as in Experiment 1. One hundred and ninety-two subjects took part to the experiment. The task in the practice and the experiment stages was a lexical decision task.

Results show that the lexical decision task would not be adapted to the study of high level understanding.

Are you Trying to be Funny? The Importance of Exaggeration on Discourse Goal Clarity.

C. O. Stewart & R. J. Kreuz

The University of Memphis (U.S.A.)

costewrt@mocha.memphis.edu

Exaggeration, or hyperbole, is one of the most common forms of nonliteral language used in contemporary American literature, second only to metaphor (Kreuz, Roberts, Johnson, & Bertus, 1996). Exaggeration also occurs about once every 200 words in undergraduates, descriptions of emotions (Link, 1999). Despite its frequency of use, exaggeration has only rarely been studied empirically (Gibbs, 1994; Kreuz & Roberts, 1993), and of the few studies that exist, many examine exaggeration only in conjunction with other forms of nonliteral language, especially irony (e. g., Colston, 1997; Gibbs, 2000; Kreuz & Roberts, 1995). Given the attention that other forms of nonliteral language, such as irony and metaphor, have received in the psychological literature, it is unfortunate that exaggeration has thus far gone largely unstudied. This paper briefly summarizes some of the existing research regarding exaggeration, presents our current findings on the effect of exaggeration on perceived humorous intent, and discusses some possibly fruitful lines of future research regarding exaggeration.

One important issue that has been addressed is why exaggeration is used. Three discourse goals have been identified for exaggeration: To clarify, to emphasize, and to be humorous (Roberts & Kreuz, 1994). Additionally, in conversation, instances of exaggeration are often perceived to be mocking (76%) and to be humorous (74%; Gibbs, 2000). Exaggeration has also been shown to be a cue for irony (Kreuz & Roberts, 1995) and for surprise (Colston, 1997). Another important issue that has been addressed is how differing degrees of exaggeration are perceived. Previous research (Colston & Keller, 1998) indicates that "realistic" exaggeration indicates as much surprise as "possible but improbable" and "impossible" exaggeration. However, ease of determining that a speaker is surprised increases as level of exaggeration increases. This research supports "The More the Better" hypothesis proposed by Kreuz, Kassler, and Coppenrath (1998). This hypothesis posits that "as the degree of exaggeration increases, the perceived appropriateness or effectiveness of a statement may also increase" (Kreuz et al., 1998, p. 97). The present research tests this hypothesis as it relates to another discourse goal known to be associated with exaggeration-namely, humor (Gibbs, 2000; Roberts & Kreuz, 1994). We predicted that higher levels of exaggeration would clarify the discourse goal of humor.

Experiment 1

Method

The materials for this study were 27 six-sentence scenarios portraying a conversation between two or more characters. The final statement in each conversation contained a reference to some numeric quantity and was potentially humorous, such as "I feel like I could sleep for 100 hours" uttered after a hard day's work. Pre-testing of the potential humor led to the elimination of six experimental texts. The remaining 21 texts were each assigned three levels of exaggeration: Not exaggerated-the most likely literal number for the topic of the statement (e.g., 10 hours in the example above); Somewhat exaggerated-the highest possible literal number for the topic of the statement (24 hours); And highly exaggerated-an impossible and very extreme number for the topic of the statement (100 hours). Additional pre-testing revealed that the assignments conformed to subjects' intuitions about level of exaggeration. In the first experiment, 30 subjects read a total of 28 scenarios (the 21 experimental texts and 7 unexaggerated filler texts). The scenarios were counterbalanced so that only one level of each experimental text was seen by each subject. For each story, the subjects then answered "How likely is it that the final speaker is trying to be funny?" on a six-point Likert scale with the endpoints labeled "not at all likely" and "very likely."

Results

The results indicate that as exaggeration increases, the perceived likelihood that the final speaker is intending to be humorous increases [$F(2, 58) = 10.08$, $p < .001$, eta squared = .26]. Significant pairwise differences were found between highly exaggerated statements (M = 4.35, SD = 1.62) and both somewhat exaggerated statements (M = 3.70, SD = 1.80, $p < .01$) and not exaggerated statements (M = 3.33, SD = 1.79, $p < .001$). However, while the means were in the predicted direction, the difference between somewhat exaggerated and not exaggerated was not significant ($p = .20$).

Experiment 2

Method

The second experiment investigated the role of prototypicality of exaggeration in addition to degree. Intuitively, it seems that by further flouting Grice's (1975) conversational maxims, using very specific (i. e., non-prototypical) values in exaggeration may convey more humorous intent than using more general (i. e., prototypical) values. The materials were 18 of the scenarios used in the first experiment and six unexaggerated filler texts (three scenarios were eliminated because the non-prototypical values were not appropriate in the given context). For the prototypicality condition, arbitrary, non-prototypical values were created by multiplying the original numbers by .91 (so 100 hours would be 91 hours). The procedure was the same as in the first experiment.

Results

Results from 30 subjects revealed neither a main effect for prototypicality ($F < 1$, ns) nor an interaction between prototypicality and degree of exaggeration ($F < 1$, ns). However, the main effect for degree of exaggeration closely replicated the findings of the first experiment [$F(2, 58) = 9.30$, $p < .001$, eta squared = .24]. Again, the highly exaggerated statements (M = 4.39, SD = 1.41) were perceived as having significantly more humorous intent than either somewhat exaggerated statements (M = 3.78, SD = 1.72, $p < .01$) or not exaggerated statements (M = 3.65, SD = 1.69, $p < .001$). There was no significant difference between the somewhat and not exaggerated statements ($p = .37$).

Discussion

The results of these two experiments indicate that high levels of exaggeration more clearly convey the discourse goal of humor than less extreme exaggeration. These findings, in conjunction with those of Colston and Keller (1998), lend credence to "The More the Better" hypothesis proposed by Kreuz et al. (1998). However, unlike Colston and Keller, the present research suggests that while there is a difference between physically possible and physically impossible exaggeration, there is no difference between physically possible exaggeration and more literal statements. The distinction between physically possible and physically impossible exaggeration is potentially important and deserves further study.

Another possibly important distinction is between discourse clarity and discourse effectiveness. Other data that we have collected using the same materials indicate that, while degree of exaggeration affects the perceived degree of humorous intent, it does not affect the perceived

degree of humor. Colston and Keller (1998) report similar findings regarding surprise. Perhaps increasing degrees of exaggeration make a speaker's discourse goals more clear but not necessarily more effective.

References

Colston, H. L., & Keller, S. B. (1998). You'll never believe this: Irony and hyperbole in expressing surprise. Journal of Psycholinguistic Research, 27, 499-513.

Gibbs, R. W. Jr. (2000). Irony in talk among friends. Metaphor and Symbol, 15, 5-27.

Kreuz, R. J., Kassler, M. A., & Coppenrath, L. (1998). The use of exaggeration in discourse: Cognitive and social facets. In S. R. Fussell and R. J. Kreuz (Eds.), Social and cognitive approaches to interpersonal communication (pp. 91-111). Mahwah, NJ: Lawrence Erlbaum Associates.

Roberts, R. M., & Kreuz, R. J. (1994). Why do people use figurative language? lang="FR" style="font-style:normal;mso-bidi-font-style:italic">Psychological Science, 5, 159-163.

Interactions on the Web Between Reviewers and Authors: Towards Reviewing an Article for the e-Journal "JIME"

A. Chevalier, N. Bonnardel & A. Piolat

University of Provence (France)

Nathalie.Bonnardel@newsup.univ-mrs.fr

The development of the Internet and the possibility to present documents on the World Wide Web (Web) is modifying not only the task of reading scientific papers, but also the design and writing of these documents (Piolat & Pélissier, 1998; Bonnardel, 1995; Bonnardel & Sumner, 1996) as well as interactions between authors of multimedia documents and readers of such documents. The research we are going to present is based on the Digital Document Discourse Environment (D3E) developed at the Open University by Sumner & Buckingham Shum (1998a). This environment assists authors and/or editors of interactive journals for the Web in embedding new media in their documents, such as hyperlinks, interactive demonstrations, and discourse facilities. This environment is especially used to publish an electronic journal (e-journal) on the web: JIME (Journal of Interactive Media in Education - www-jime.open.ac.uk). It aims also at promoting interactions between authors, reviewers and readers of scientific papers presented on the Web. The use of such an environment should profoundly affect processes of writing and reading articles as well as the process of paper review, since it facilitates exchanges between different stakeholders or partners during the review process. Therefore, the review process induced by the environment JIME should differ from traditional review process, which is basically non interactive: Comments from the reviewers are received afterwards by the author and aim at inciting him or her to improve the submitted paper. Moreover, the review is usually anonymous (Cicchetti, 1991), which is not the case with the JIME environment (Sumner, Buckimgham Shum, Wright, Bonnardel, Piolat & Chevalier, in press).

The editors of the JIME e-journal wished to reduce the time of review and to open the review process not only to reviewers and authors but also to readers. Indeed, different stakeholders (reviewers, authors, readers and editors) can interact throughout a new review process. The

submitted paper is assessed with regard to criteria specified by the JIME e-journal (e.g., quality of writing, originality of ideas,...) and according to each paper's own structure (Sumner & Buckingham Shum, 1998b):

1. There is first a one month 'closed review' period, during which reviewers (usually three) and authors discuss and debate the article on a document-centered discourse interface presented in a secret site. Towards this end, reviewers write their comments into a specific window of the interface, which is directly connected (through hyperlinks) to the different parts of the article reviewers wish to comment on.

2. Then, the secret site containing the submitted article and review debate is made available to the public for a one month 'open review' period. During this period, readers are able to join the discussion and the reviewers can pursue their comments. In order to know more about this new review process, we are going to study the mode of interactions, which occur between authors who submit papers, reviewers and, possibly, readers. In order to better know the real effect of such environments, two main questions need to be answered: (1) Is the dialogical environment offered by the e-journal JIME profoundly affecting the review process? (2) What is the extent of incitements to modify the text ?

Method

Procedure. We randomly chose one of 19 available papers ("Integrating interactive media in courses") for this analysis. Comments from the different stakeholders participating in the review process (authors, reviewers, readers and editors) were filed as soon as they appeared on the Web site of JIME. The chronical of email messages includes the transmitter's name, the title of the message and the body of the message. Each message can have from 9 words to more than 400.

Several quantitative analyses were performed on this corpus, in order to: (a) identify messages associated to answers (whatever the delay of answer) as well as messages which were associated to no reaction; (b) determine linguistic indicators, which show that stakeholders pay attention to previously received messages (lexical repeat, insertion of "copy-paste" in the answer, etc.); (c) determine whether all the suggested evaluative categories (5 criteria and, for the selected paper, 12 chapters) were used during the review process; (d) determine the extent of incitements to modify the text.

Results

We observe that only the second reviewer (R2) interacts with the author (A). R1 and R3 do not interact with A, nor between themselves. We also notice that A does not react to readers' messages. Despite this lack of noticeable dialogue, the different stakeholders try to collaborate through verbal indications. In the titles of messages, reviewers mainly use the interrogative clause to promote a dialogue (R1=100%, R2=55% et R3=50%). This is not the case of the author (A=0%), who uses the affirmative clause. Moreover, lexical items included in the titles of previous messages are reused in titles of new messages in order to reinforce the dialogical consistency. In the body of messages, only the editors and one reader insert (in 5 messages) the first name of the other stakeholders in order to explicitly refer to previously expressed ideas. The reviewers do not use this procedure, except R2 (who directly asks for the other stakeholders' points of view). The use of certains articles (I/you - my/your...) shows another type of collaboration, which appears only during the dialogue R2-A. Finally, the insertion of extracts of previous messages through "copy-paste" allows the reference to ideas already expressed by other stakeholders. This procedure was

used only two times by the author and in one message. Inthe analysis of that content of the exchanged messaged, we notice that:

- the dyade R2-A develops two conversational nucleus, which refer to only 17,6% of the evaluative categories proposed by JIME (5 criteria and 12 chapters);

- the messages sent by R3 refer to 64,7% of the evaluative categories;

- the messages sent by the readers are related to the main evaluative categories used by the reviewers.

The reviewers and the readers do not try massively to force the author to modify her text. There is only a debate of ideas (only 32% of the messages request a modification of the text). The author does not answer to all the requests: she only takes part to the debate of ideas, especially while interacting with R2. No request from the reviewers or the readers aims at improving the quality of the writing. The required modifications of the text strictly aim at adding information elements (such as experimental data already published but not reported in the submitted paper).

Conclusion

These first results do not allow us to conclude to a massive modification of the review process. However, a collaborative process appears through noticeable thematic repetitions (choice of titles of the messages, etc.). Moreover, the goal of the reviewers and readers was not to incite the author to modify her text. To summarize, we observe no cooperative modification of the text, but a collaborative incitement to changes. Therefore, an important step is still needed to promote a really open and dialogical review process.

References

Bonnardel, N. (1995), L'évaluation dans les activités de conception : Vers une assistance ergonomique. Psychologie Française, 40(1), 99-110.

Bonnardel, N. & Sumner, T. (1996). Supporting evaluation in design. Acta Psychologica, 91, 221-244.

Piolat, A. & Pélissier, A. (1998). Etude de la rédaction de textes: contraintes théoriques et méthodes de recherches. In A. Piolat & A. Pélissier (Eds.), La rédaction de textes. Approche cognitive, (pp. 225-269). Lausanne: Delachaux & Niestlé.

The Impact of Verbal Information on the Aesthetic Experiences to Visual Art

K. Millis, K. Schaefer & B. Salmon

Northern Illinois University (U.S.A.)

TJOKKM1@wpo.cso.niu.edu

One important way in which discourse is communicated is via visual art (e.g., illustrations, paintings, and photographs). Similar to the textbase and situation models studied by discourse psychologists, when viewers comprehend artwork, they construct a representation of the explicitly shown concepts and a representation of how they conceptually inter-relate to one another along with the artist's intentions. (Solso, 1994).

We focused on one important aspect of comprehending visual art, namely the viewer's aesthetic experience. Aesthetic experiences include liking, interest, inferences, emotions, and understanding. We were interested in whether and how verbal information in the form of the title influences the aesthetic experience. This question is related to both psychology and art communities. In psychology, it is related to the issue of how verbal information interacts with basic mechanisms involved in perceiving visual information. One might expect that verbal information would have little or no impact on perceiving visual art since verbal and visual information are processed in different areas of the brain. However, verbal information might influence the comprehension and appreciation of visual art since both words and images would presumably trigger concepts in working memory. Within Kintsch's construction-integration framework (1988), the words and images would activate their respective concepts in working memory in a bottom-up fashion (construction) and then the activation levels of the concepts would settle into a stable configuration as they are integrated into a representation (integration). Therefore, if the verbal information adds information to what can be readily inferred from viewing the artwork, the words would affect the resulting representation, and possibly the aesthetic experience.

Within the art community, the question of whether titles influence the aesthetic experience is founded in two schools of thought. According to Formalism, the quality of a piece of artwork is independent of its context. The title serves as label for the artwork and nothing else. Therefore, the contents of a title should have no impact on its evaluation. However, according to Constructivism, context may have an impact on aesthetics. Because an artwork is viewed in the context of its title, titles could have an impact on the viewer's appreciation of the art, and the accompanying aesthetic responses.

Of course, the influence of titles is well known on comprehending ambiguous text (Bransford & Johnson, 1972) and pictures (Dooling & Lachman, 19971). However, there is considerable less known about the influence of verbal information on understanding and appreciating real visual art. A study by Cupchik, Shereck, and Spiegel (1994) measured the aesthetic responses to sculptures when they were paired with different types of textual information. Like Franklin et al. (1993), they showed the artworks twice, but the first exposure the artwork was shown without textual information, and once with textual information. They found that the effect of textual information depended on the artist, type of textual information, and the gender of the participant. One criticism of this and similar research (e.g., Franklin, Bechlen, & Doyle, 1993) is that participants saw the same piece of artwork with titles or with and without text. This raises the issue of whether demand characteristics affected participants' responses.

Method

Materials. Thirty artworks were used in the study. One-half of artworks were representational colored illustrations; the other half were photographs. Although they were not known to our participants, they had been published in art books.

Titles. Participants saw each artwork accompanied by a title. In a No-title condition, the artwork was labeled untitled. In a Descriptive-title condition, the title was a simple declarative sentence that described the scene the art depicted. For example, the Descriptive title for one illustration that depicted a maid walking down a staircase carrying some folded sheets was "A Maid Carrying Sheets Walking Down a Staircase". In an Elaborative-title condition, the title suggested an

explanation or one possible interpretation of the scene. The Elaborative-title for the maid artwork was Sickness. Title condition was a within-subjects factor, so that each participant saw 10 examples of No-title, Descriptive-Title, and Elaborative-Title artworks. One-half of the participants were asked to talk aloud as they inspected the artwork. The verbal protocols were later transcribed and coded.

Procedure. Participants were given a booklet with the 30 artworks and their respective titles. For each artwork, they wrote down its title (to ensure that they would encode the title), inspected the artwork, talked aloud if they were in the talk condition, and answered questions assessing their aesthetic responses. Using 6-point Likert-type scales, they answered questions regarding their understanding, interest, emotion, enjoyment, and number of thoughts elicited by artwork (1 = no enjoyment, 6 = great enjoyment).

Results

The ratings for each measure (understanding, interest, emotion, enjoyment, and thoughts) were moderately correlated (r = s > .60). In order to simplify the presentation here, we collapsed the ratings into a single appreciation score.

The means are presented below:

	No title	Descriptive	Elaborative
No talk	3.56	3.61	3.77
Talk	3.32	3.37	3.58

We found the following statistically significant pattern (all ps < .05 on both subject and item analyses): Elaborative > Descriptive = No title, indicating that the title affected the aesthetic responses. In particular, when the title afforded an interpretation which would not normally be generated (i.e., the elaborative condition), aesthetic responses increased. Surprisingly, having talked aloud decreased responses. However, the increase from the Descriptive to the Elaboration condition was larger in the talk condition than in the no talk condition. This finding suggests that talking aloud might have engendered inference-making given there were multiple interpretations (as indicated by the elaborative titles).

Other findings: The verbal protocols indicated that aesthetic responses increased to the extent that the viewer had generated inferences about the story world. In Experiment 2, we found that random titles did not increase aesthetic responses, suggesting that aesthetic responses are only increased if the viewer generated a coherent representation of the art/title. In Experiment 3, we found that instructions to ignore the title because they were made up did not alter the pattern, suggesting concepts, once activated in working memory, is subject to integration.

Overall, the data from three experiments suggest that aesthetic responses increase when multiple concepts get triggered while viewing artwork, but only when they can be integrated into a coherent representation.

References

Bransford, J.D. & Johnson, M.K. (1972). Contextual prerequisites for understanding: Some investigations of comrpehension and recall. Journal of Verbal and Verbal Behavior, 11, 717-726.

Cupchik, G.C., Shereck, L., & Spiegel, S. (1994). The effects of textual information on artistic communication. Visual Arts Research, 62-78.

Dooling, D. J. & Lachman, R. (1971). Effects of comprehension of retention of prose. Journal of Experimental Psychology, 88, 216-222.

Franklin, M.B., Bechlen, R.C., & Doyle, C.L. (1993). The influence of titles on how paintings are seen. Leonardo, 26, 103-108.

Solso, R. L. (1994). Cognition and the visual arts. Cambridge, MA: The MIT press.

Conjoined Effects of Long-Term Working Memory Representation, Text Difficulty and Reading Purpose on Reading Strategies in Text Problem Comprehension

M. A. Schelstraete & I. Walschaerts

University catholic of Louvain (Belgium)

schelstraete@exco.ucl.ac.be

The point of departure of the present research was an experiment published by Mills, Diehl, Birkmire & Mou (1995). In their experiment, participants had to read two procedural texts twice in a self-paced sentence-by-sentence reading task. Before reading the first text, participants were said that they had to read the texts either in order to recall them (Read-to-Recall condition) or in order to perform the task described by the text (Read-to-Do condition). However, after the second reading of the first text, both groups were informed that they had actually first to recall the text and second to carry out the task. Three dependent variables were analyzed: Reading times, number of recalled propositions and number of correct actions. Results showed that purpose for reading influenced participants' performances: Read-to-Do participants performed the task better and Read-to-Recall participants recalled the text better. Read-to-Do participants recalled less of the information judged to be less important for performing the task (hereafter irrelevant elements) than Read-to-Recall participants. Reading times varied as a function of the importance of the information to task performance, differed for the two texts, and was faster for the second reading. The authors interpreted their results as showing that participants in the Read-to-Recall condition mainly used their propositional representation while participants in the Read-to-Do condition mainly used the situation model.

The present experiment used a similar task to the Mills et al.'s research but with some important modifications designed first to solve some methodological problems observed in this previous experiment and second to test the long-term working memory (LT-WM) approach proposed by Ericsson and Kintsch (1995). First, the texts contained more irrelevant elements for solving the task: Relevant information was embedded in a cover story where it represented about 20% of the

propositions for one text (the side-car story) and about 50% for the other one (the tea ceremony story). Second, the texts did not describe a task to perform but a problem to solve, which was intended to allow a better control of the propositional representation and of the situation model built during reading. The situation model was supposed to correspond to the problem space described by the story while the propositional representation would match the propositional content of the story. Thirdly, participants performed only one task: Either Read-to-Recall or Read-to-Do, which would allow to have more homogenous patterns on both texts and to exploit the variable of text difficulty (cf. the following point). Fourthly, the text difficulty was manipulated so that the problem space was more difficult to build in one text ("The tea ceremony", an isomorphic problem of the Tower of Hanoi) than in the other one ("The side-car problem" with a cross-river space problem).

It was expected that, in the first reading, the readers would built different retrieval structures in LT-WM as a function of their reading purpose, which would appear mainly in the pattern of reading times during the second reading of the text: (a) When the participants had to read the text in order to solve the problem described in it, it was hypothesized that the retrieval structures would be used during the second reading of the text to select more quickly and more precisely the relevant elements to achieve the task. So, in the Read-to-Do condition, the reading times of the relevant elements would be lower relative to the reading times of the irrelevant elements during the second reading of the text in comparison with the first reading; (b) When the participants had to recall the story, the effect of the relevance factor in the building and use of the retrieval structures was expected to be less important since the irrelevant elements were assumed to be more important for the recall of the story than for the problem resolution.

These effects were expected to be modulated by the level of text difficulty as the problem space was expected to be more difficult to build for the Tea Ceremony story. More precisely, the retrieval structures were expected to contain relevant elements for the task to perform (recall or solving). In the solving condition (Read-to-Do), only some elements of the text were relevant whereas in the recall condition (Read-to-Recall), all elements were relevant since participants had to recall all the text (Note, in this presentation, the distinction was maintained for both conditions).

Method

Thirty-six French-speaking students in Psychology at the Université catholique de Louvain participated in the experiment. All were second-grade students who received a credit course for their participation in the experiment. Half of the students were randomly assigned to the Read-to-Recall condition and the other half, in the Read-to-Do one. In both conditions, the order of the texts was counterbalanced. In addition to the two experimental texts, the participants read (1) a text designed to record their mean reading rate, (2) short texts designed to detect eventual reading comprehension difficulties, (3) a training text presented before the experimental texts. All tasks were presented on a computer. Except for the comprehension test where the time presentation was fixed (only the answers were recorded), the other tasks were presented in a sentence-by-sentence self-paced reading task; For the experimental texts, both reading times and verbal protocols were recorded.

Results

In this paper, we mainly focused on the analyses of reading times of the experimental texts in order to show how retrieval structures built during a first reading of a text were used during a

second reading. Reading time data were analyzed by dividing the reading times for relevant elements by the reading times for irrelevant elements in each text. Results showed that this proportion was higher in the Read-to-Do condition for the second reading and particularly for the more difficult text. These results showed that, with regard to the first reading, reading times of relevant elements increased only in the Read-to-Do condition and especially for the text whose the problem space was more difficult to build. In the terms of the LT-WM hypothesis, the results nicely match the idea that skilled readers encode quite large retrieval structures that are maintained accessible for subsequent processing. In addition, they show that the nature of such retrieval structures varies according to the reading purpose and to the level of text difficulty. It could be supposed that, when the participants read a text in order to solve the problem described by the text, the retrieval structures would be nearer to the situation model which matches in this case the problem space; When they read a text in order to recall it, the nature of the retrieval structures would be more propositional, nearer to the text base.

References

Ericsson, K.S. & Kintsch, W. (1995). Long-term working memory. Psychological Review, 102, 211-245 .

Mills, C.B., Diehl, V.A., Birkmire, D.P., & Mou, L.C. (1995). Reading procedural textes : Effects of purpose for reading and predictions of reading comprehension. Discourse Processes, 20, 79-107.

The Effect of Verbal Context on Picture Recognition: Initial Support for Perceptual Symbol Theory

R. A. Stanfield & R. A. Zwaan

Florida State University (USA)

stanfiel@darwin.psy.fsu.edu

The purpose of the current experiment was to examine predictions generated by two competing theories of mental representation, amodal and perceptual symbol systems. Whereas perceptual symbol systems predict an analogous relationship between the symbol and its referent, amodal systems predict an arbitrary relationship between the symbol and its referent (Barsalou, 1999). Perceptual and amodal symbol systems make different predictions about language comprehension. For example, perceptual but not amodal symbol systems predict that comprehenders commit to a specific representation of an entity. For instance, if the discourse implies that a referent has a specific spatial orientation, then a perceptual symbol should represent this orientation whereas an amodal representation should not.

In the present experiment, participants read sentences that implied that an object was either in a vertical or a horizontal orientation (e.g. "The carpenter pounded the nail into the floor," "The carpenter pounded the nail into the wall"). They then viewed a picture of the object (e.g. nail) in either a vertical or horizontal orientation. The combination of sentence and picture produced two conditions (Match versus Mismatch). The prediction derived from Perceptual Symbol Theory is that

recognition latencies to objects in the Match condition should be shorter than recognition latencies to objects in the Mismatch condition, because comprehenders have committed to a specific representation of the object. If this representation matches the picture that is presented afterwards, the object should be recognized more quickly than when there is a mismatch. In contrast, Amodal Symbol Systems predict no effect of match or mismatch, because they assume that comprehenders have not committed to a specific representation.

Method

Forty undergraduates from The Florida State University participated for course credit in the experiment. One hundred and five black and white line drawings (Snodgrass & Vanderwart, 1980) were used, as well as 129 sentences (81 filler, 48 experimental). Of the 48 experimental sentences, half implied an object in a horizontal orientation and half an object in a vertical orientation. Thus, there were 24 objects described in two orientations. Participants saw only one version of each sentence. The Flags Test (Thurstone & Jeffreys, 1956) was used to measure participant's spatial ability. Participants were classified as either high ability or low ability based on a median split of the Flags Test scores.

Participants were instructed to read each sentence for comprehension. After reading the sentence a picture appeared in the center of the screen. Participants responded "yes" or "no" by pressing one of two labeled keys to indicate if the object had been mentioned in the sentence. Half of the trials required a "yes" response and half a "no" response. In addition, approximately half of the filler items required the participants to retype the sentence after making the verification decision. This task was included to ensure that participants were carefully reading each sentence. Each picture was presented both horizontally and vertically and picture orientation was crossed with Sentence Type ("vertical" versus "horizontal"). A pilot study demonstrated that the pictures used in this experiment did not have a recognition bias for either orientation, nor were vertical sentences more likely to be matched with their associated picture than were horizontal sentences, or vice versa. After finishing the sentence-picture verification task participants completed the Flags Test. The entire procedure lasted about 45 minutes.

Results

The mean recognition latencies were 838 ms (SD = 331) for the Match condition versus 882 ms (SD = 329) for the Mismatch condition. A 2 (Orientation) X 2 (Match) X 2 (Spatial Ability) X 4 (List) analysis of variance (ANOVA) revealed a main effect of Match [$F(1, 32) = 5.36, p < .05$], but no effect for Orientation [$F < 1$], and no significant interaction between Match and Orientation [$F(1, 32) = 1.57$].

The Flags Test (M = 72; range = 21 - 118) did correlate significantly with both the Match and the Mismatch conditions ($r = -0.33; p < .05; r = -0.38; p = .01$ respectively), suggesting that individuals with better spatial ability responded faster than individuals with poorer spatial ability. The Flags Test did not interact with Match [$F < 1$] or Orientation [$F(1, 32) = 1.45$], nor was there a significant Flags Test X Match X Orientation effect [$F < 1$]. Given the Flags Test's lack of participation in the analyses of interest, it was excluded from further discussion.

Discussion

To our knowledge, this is one of the first studies that has obtained direct empirical evidence in support of Perceptual Symbol Systems. An important caveat, however, is that even though amodal systems would have predicted a null effect in the current experiment, such systems could still explain the results in a post-hoc manner, e.g., by assuming that readers construct propositions regarding the (implied) orientation of objects. This should not be seen as support for amodal systems, however. It has been noted that such systems are virtually unfalsifiable (e.g. Barsalou, 1999) and thus capable of explaining anything post hoc. It should also be noted that perceptual symbols provide a more parsimonious explanation of this effect than would amodal systems. As Barsalou argues, the theory that both predicts the effect and explains said effect with greater parsimony should be favored over the theory that accomplishes neither.

Future research will examine this effect in relation to situation model theory. An analogous relationship between the perceptual symbol and the referent has important consequences for our understanding of situation models and situation model construction. Understanding situation-model construction in terms of perception can open up many new avenues of research and help shed new light on old findings.

References

Barsalou, L.W. (1999). Perceptual symbol systems. Behavioral and Brain Sciences, 22, 577 - 660.

Snodgrass, J.G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. Journal of Experimental Psychology: Human Learning & Memory, 6, 174 - 215.

Thurstone, & Jeffreys (1956). Flags: A Test of Space Thinking. Chicago, IL: Chicago Industry Service.

Long-Term Working Memory: Some Empirical Results

C. Bellissens & G. Denhière

University of Provence (France)

cedrickb@newsup.univ-mrs.fr

There is a contradiction between the processes implied in text processing and the limits of short-term working memory (Kintsch, 1998). Text Comprehension is a dynamic process that implies availability of a large amount of elements and structures in Working Memory during reading. But almost all comprehension theories including a WM component imply that WM capacity is limited (Kintsch & van Dijk, 1978; Just & Carpenter, 1992). Hence, a limited capacity in WM does not permit a large number of elements to be stored at one time. As a consequence, Ericsson and Kintsch (1995) have proposed a new conception of WM during reading.

Long Term Working Memory (LT-WM). The LT-WM basic assumptions can be summarized as follows: (1) During reading, new information activates knowledge structures in LTM; (2) Activated knowledge structures can form a stabilized set, associated with encoded information; (3) Encoded

information can keep up meaning overlap relationship between propositions; and (4) activated knowledge structures associated with encoded information is a retrieval structure consisting in a stabilized set of retrieval cues (knowledge structures) associated with a text base (encoded information) in LT-WM.

Reading Interruption Procedure. According to the classic view of WM, during reading, clearing STM necessarily leads to comprehension impairment because readers should retrieve encoded information from LTM which is classically considered as a processing impairment. Glanzer, Dorfman and Kaplan, (1981) and Glanzer, Fisher and Dorfman (1984) used a reading interruption procedure consisting in presenting a text and interrupting its presentation by an interrupting task (e.g. reading another text). In all the cases they used this procedure: (1) The interruption did not have an effect on the number of correct answers to comprehension questions; (2) The interruption had a significant effect on the reading time of the first sentence of the second part of the interrupted text, increasing this critical sentence reading time of about 400 ms; and (3) When last sentence of the first part of the text was reinserted at the resumption of the interrupted text, the effect of the interruption disappeared. To explain the ininterruption effect, Glanzer and his colleagues claimed that at the resumption of the text the readers had to recreate the surface structure of the one or two sentences presented before the interruption.

Hypotheses. The interruption effect cannot only be explained by the reinstatement of the surface structure. After the interruption, readers have to retrieve encoded information from LTM where the semantic trace of textual information is stronger than the surface trace (Kintsch, Welsh, Schmalhofer & Zimny, 1990). Therefore, we hypothesized that readers reading a familiar text are able, during encoding to elaborate a retrieval structure consisting of encoded information associated with knowledge structures that act as semantic retrieval cues. The semantic retrieval cues can be reinstated in STM during reading, provide a rapid access to associated encoded information and maintain the availability of the elaborated structure.

Method

We designed 3 experiments using the reading interruption procedure. Texts we used had exactly the same structure.

Experiment 1 (Denhiere & Bellissens, 1997). Several kinds of cue sentences were inserted at the reading resumption of the interrupted text in order to compare efficiency of either surface or semantic retrieval cues.

Experiments 2 and 3 (Denhière & Bellissens, 1999). In these experiments, encoding conditions of information were manipulated. The texts were presented in two versions according to word frequency (High or Low) of sentences of the first part. In the first experiment, reading was self-paced and in the second the reading was time-limited. Comprehension questions were also manipulated: Surface questions relied on surface structure and content questions relied on meaning of the texts.

Results and Discussion

In the three experiments, the interruption had no effect on the number of correct answers to comprehension questions. The interruption increased the reading time of the first sentence of the second part of the interrupted texts. Thus, clearing ST-WM does not lead to comprehension impairment because readers are able to rapidly retrieve encoded information from LT-WM.

Experiment 1. When reading resumed, readers could rapidly reinstate cues to retrieve encoded information from LT WM, and the retrieval cues could be represented at a semantic level and not only at a surface level.

Experiment 2 and 3. At the reading resumption of the interrupted text, the retrieval times were significantly shorter when the interrupted text was familiar than when it was not familiar enough and reading a familiar text, readers mainly encode meaning of the text rather than surface structure. Therefore, during encoding of a familiar text, readers can activate knowledge structures acting as semantic retrieval cues associated with encoded information.

Conclusion: Towards a simulation of LT-WM

Results we obtained are consistent with the LT-WM theory: When reading a familiar text, readers constructs a retrieval structure consisting in encoded information associated with knowledge structures that act as semantic retrieval cues. Readers can use functional relations (Caillies, Denhiere & Kintsch, accepted) or causality/consequence (Aguilar, 1998) to construct a retrieval structure. We will describe how this kind of process has already been simulated using construction-integration model (Kintsch, 1988) in which a knowledge base was associated with a text base of a given text. We will also explain how a realistic knowledge base can be constructed using Latent Semantic Analysis model (Landauer & Dumais, 1997).

References

Aguilar, N. (1998). Inferences sur la consequence probable des evenements: production ou recuperation? Spoken paper presented at the annual meeting of the ARC Society. Paris.

Caillies, S., Denhiere G., & Kintsch, W (1998). Representation of knowledge in memory: Evidence from primed recognition. Spoken paper presented at Cogsci98 meeting. Madison.

Denhiere, G., & Bellissens, C. (1997). Retrieval from long-term working memory. Spoken paper presented at the 8th annual meeting of the Society for Text and Discourse. Madison.

Denhiere, G., & Bellissens, C. (1999). Long-Term Working Memory: A frequency effect on retrieval time during reading. Spoken paper presented at 40th annual meeting of the Psychonomics society. Los Angeles.

Ericsson, K.A., & Kintsch, W. (1995). Long-term working memory. Psychological Review, 102, 211-245.

Glanzer, M., Dorfman, D., & Kaplan, B. (1981). Short-term storage in the processing of text. Journal of Verbal Learning and Verbal Behavior, 20, 656-670.

Glanzer, M., Fischer, B., & Dorfman, D. (1984). Short-term storage in reading. Journal of Verbal Learning and Verbal Behavior, 23, 467-486.

Just, M.A., & Carpenter, P.A. (1992) A capacity theory of comprehension: Individual differences in working memory. Psychological Review, 99, 122-149.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. Psychological Review, 95, 163-182.

Kintsch, W. (1998). Comprehension: A paradigm for cognition. Cambridge. University Press.

Kintsch, W., & van Dijk, T.A. (1978). Towards a model of text comprehension and production. Psychological Review, 85, 363-394.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The Latent Semantic Analysis theory of the acquisition, induction, and representation of knowledge. Psychological Review, 104, 211-240.

Perspective Effects on Fixation Times and Memory for Text

J. Kaakinen*, J. Hyönä* & J. M. Keenan**

***University of Turku (Turkey)**

****University of Denver (U.S.A.)**

johkaa@utu.fi

A well-known study in cognitive psychology is Anderson and Pichert's (1978) demonstration that it is the perspective active during retrieval that determines what is recalled from a text, and that the perspective active at encoding appears to have little effect on memory. Baillet and Keenan (1986), however, showed that Anderson and Pichert's surprising finding was due to certain methodological factors. While they also found an effect of the recall perspective, it was the encoding perspective that ultimately constrained what could be recalled. Thus, the new perspective given at recall did not increase the amount of material recalled; it only affected the omission of information: Information not cued by the recall perspective was more likely to be omitted than information relevant to the recall perspective. In all, Baillet and Keenan showed that although the retrieval perspective is capable of making some information more accessible for output, memory for text is primarily constrained by the accessibility of information determined by the encoding perspective. In Experiment 1, we aimed to get further support to this view by examining the effects of reading and recall perspectives on text memory. However, questions remain as to exactly how the encoding perspective exerts its influence. Clearly, it gives readers a guide for determining what information is important and thus which information to attend to. But what does it mean to give increased attention? Probably the simplest way to view increased attention is to assume that it means increased processing time. According to this encoding time hypothesis, perspective relevant information is processed longer than irrelevant information and thus is better remembered. In Experiment 2, we tested the encoding time hypothesis by using the eyetracking technique to record readers' fixation times. We expected to find longer fixation times for perspective-relevant than perspective-irrelevant information.

Moreover, we were interested whether readers, working memory span (WMS) modulates the possible perspective effects on text memory and text processing. A study by Lee-Sammons and Whitney (1991) suggests that perspective effects in text recall. Lee-Sammons and Whitney divided capacity differences may modulate the extent to which readers show participants into low-, medium-, and high-span groups on the basis of a WMS test. The results showed that low-span participants recalled more information relevant to the encoding perspective than relevant to the

retrieval perspective, whereas high-span readers recalled equal amounts of relevant and irrelevant material. In sum, as the WMS decreased, the effect of the encoding perspective was found to increase. Along the lines of the capacity theory of comprehension (Just & Carpenter, 1992), Lee-Sammons and Whitney proposed that readers with a high working memory capacity have enough resources available to maintain all the text information, whereas readers with a low working memory capacity have to compensate for their limited resources by selectively maintaining only the most pertinent information. Experiment 1 examined the individual differences in the perspective effects on text memory. If working memory capacity constrains encoding of text, as suggested by Lee-Sammons and Whitney (1991), the low-span readers should show a stronger effect of the encoding perspective on text recall than high-span readers. In Experiment 2, the individual differences in the perspective effects were further examined for on-line text processing.

Experiment 1

In Experiment 1, 32 participants were given a perspective from which to read the Housewalk text adopted from Baillet and Keenan (1986). Recall was collected both immediately with the same perspective and after a one week delay with a new perspective given as a recall cue. Readers, WMS was assessed by the reading span test (Daneman & Carpenter, 1980). In the immediate recall, participants recalled more reading perspective relevant than reading perspective irrelevant material. After the perspective shift in the second recall, there was no increase in previously unrecalled information. Readers recalled more reading perspective relevant than reading perspective irrelevant material, even though they omitted more material that was irrelevant to the recall perspective than relevant to the recall perspective. Thus, the results support the view that it is the encoding perspective active during reading that constrains text memory. Contrary to Lee-Sammons and Whitney (1991), readers with relatively high WMS showed a greater effect of the encoding perspective than low-span readers. This result implies that the correlation between reading span and perspective effect may be more due to differences in strategic processing rather than (or in addition to) differences in WM capacity (see below).

Experiment 2

In Experiment 2, 64 participants read an expository text from a given perspective while their eye-fixation patterns were recorded. The text described four relatively unknown small countries. A complete recall of the text was collected and readers' WMS was tested. Four processing measures were derived from readers' eye fixation patterns, (1) first-pass fixation time, (2) forward fixation time and (3) reinspection time during first-pass reading, and (4) look-back time. All fixations landing on a target segment during the initial reading of the segment are defined as first-pass fixations, whereas fixations returning back to a target segment from a subsequent segment are called look-back fixations. First-pass fixations were further divided into forward fixations and reinspective fixations; forward fixations land on an unread part of the target segment, while reinspections consist of fixations returning to an already processed part of the target segment. Perspective-relevant text segments attracted longer fixation times than perspective-irrelevant text segments and perspective-relevant information was also better recalled than perspective-irrelevant information. Thus, the results imply that increased processing time for perspective-relevant information is related to increased memory for perspective-relevant information, and support the encoding time hypothesis. Readers with relatively high WMS showed a greater effect of the reading perspective on first-pass fixation times than low-span readers. The high-span readers seemed to use the reading perspective already during the initial phase of processing: They recognized the perspective-relevant information quickly, paid extra attention to relevant information and went quickly through the irrelevant information with relatively few forward fixations. Moreover, they reinspected and later went back to the perspective-relevant text segments more than to perspective-irrelevant text segments. Medium- and low-span readers were not selective during the

initial processing (i.e., forward fixations), but showed a more delayed effect: They reinspected the perspective-relevant information longer than perspective-irrelevant information. Low-span readers also went back to perspective-relevant text segments more than to perspective-irrelevant text segments. These results suggest that high-span readers employ a selective reading strategy (Anderson, 1982; Reynolds, 1992) in the initial phase of processing when given a reading perspective, and give further support for the view that the reading span test taps strategic processing.

Conclusion

The results of Experiment 1 confirm those obtained by Baillet and Keenan (1986): The perspective active during reading constrained what was recalled of text. Experiment 2 showed that increased memory for perspective-relevant information was related to increased processing time for perspective-relevant information, thus supporting the encoding time hypothesis. The results also suggest that high-span readers employ the selective reading strategy (Anderson, 1982; Reynolds, 1992) in the initial phase of processing when given a reading perspective. Moreover, the results support the view that the reading span task may actually tap something else than just available resources for processing, namely, strategic allocation of attention (Carpenter & Just, 1989). Carpenter and Just proposed that the individual differences observed in the reading span test may in part reflect a more efficient or intelligent processing strategy. They registered eye movements during a standard reading span task and observed high-span readers to spend more time in gazing at the to-be-remembered sentence-final words than low-span readers, but less time reading the sentence up to the final word, particularly as memory load increased. In other words, high-span readers strategically allocated relatively more time to task-relevant text elements and processed more superficially less relevant text elements - a finding completely in line with what we observed here for text processing.

The Effects of Testing and Delay on Recall and Learning of Information from Texts

S. Gulgoz, M. E. Aktunc, D. Senol & H. A. Van

Koc University (Turkey)

sgulgoz@ku.edu.tr

Research in cognitive psychology has revealed many times that learning from texts could be significantly facilitated by devising texts that are more coherent in structure. Coherence of a text is defined as the degree of interconnectedness among the arguments in the text. In a sufficiently coherent text there are a reasonable number of connections between the arguments and these connections enable the reader to make the inferences necessary to reach a full comprehension of the text content. However, some texts lack some of these connections and this deficiency has negative effects on the comprehension of these texts. Specifically, because the text does not provide some of the connections between propositions, the information is not conveyed with sufficient clarity for the reader to comprehend the information in the text completely. In several studies on this issue, the necessary connections that were lacking in the original text were inserted in the text and a more coherent version of the text was obtained. Previous research has repeatedly

shown that participants who read the coherent version of a text comprehended and recalled text information better than the participants who read the original version.

The theoretical basis of the above findings was provided by Kintsch and his colleagues (1988, 1994, 1998; Kintsch & vanDijk, 1978; Miller &Kintsch, 1980). According to this theory, besides the superficial level of the text, there are two levels in the processing of a text. The first level is called "textbase" and it is the level in which the linguistic content of the text is read and understood. In other words, only the content that the text reveals is comprehended. The second level is called "situation model" and it refers to a deeper level of understanding. At this level, relations between the content of the text and the content of the reader's long term memory are formed. The reader reads the text, understands what is conveyed in it, finds the information already present in his or her long term memory that relates to the information conveyed by the text, and forms relations between the present information in long term memory and the information in the text. In this way, a permanent representation of the text, the "situation model", is formed in long term memory. One type of hypothesis derived from Kintsch and his colleagues' model is that revising texts to make them more coherent would facilitate comprehension and learning. This hypothesis was confirmed in several studies in the last decade as well as qualifications by Kintsch and his colleagues (e.g., Kintsch, 1994; McNamara & Kintsch, 1996). They argued that although coherence of the text may prove beneficial for the recall of the text and by the readers with little knowledge, readers with sufficient knowledge would benefit from the gaps because they would compel integration with prior knowledge.

In addition to research that conveys contradictory findings regarding these arguments (e.g., Gulgoz, Kumkale, Aktunc, Eskenazi, in press; McNamara & Kintsch, 1996; McNamara, Kintsch, Songer, & Kintsch, 1996), one can search for the initiating mechanisms of extra processing necessary for the integration of text information with prior knowledge at various levels. One source has been observed to be needed for cognition (Gulgoz, in press; Gulgoz, Kumkale, Aktunc, Eskenazi, in press). Need for cognition is defined as an individual difference in readiness and willingness to engage in cognitively demanding activities such as thinking, reading, problem solving etc. Research shows that people who score high on the need for cognition scale also achieve higher scores on tests of comprehension and recall (Cacioppo, Petty, Feinstein, & Jarvis, 1996; Gulgoz & Sadowski, 1995; Sadowski & Gulgoz, 1992; Gulgoz, in press). Another source of initiating cognitive activity that would give way to the construction of a situation model is asking questions about the text. These questions would be the kind that required the integration of prior knowledge with the text information.

In the present study, in addition to the effects of answering questions, the effects of text coherence and need for cognition were studied. Another manipulation was done about the time of tests of comprehension and recall. It has been the dominant paradigm in learning from text research to give immediate recall and comprehension tests soon after the participants read the text. However, it is our belief that this paradigm is not appropriate to test the effects of the variables on long term learning. Therefore, immediate and delayed tests of comprehension and recall were given in the present study. The study had two phases that were one week apart. In the first phase, half the participants were given the original version of a text about tool using characteristics of the crow which was around 800 words long. The other half were given a version that was revised to be more coherent by the researchers. Half the participants were given only a free recall protocol and the other half were given both the free recall protocol and a test consisting of 12 short answer questions. The free recall protocol tested the recall and the 12-question comprehension test of inference questions aimed to measure the construction of the situation model. In the second phase, all the participants were given both recall and comprehension tests. In addition, they were given a scale measuring need for cognition developed by Cacioppo and Petty (1982).

There were four major analyses that were conducted on the data. The first analysis was on free recall data and examined the effects of delay, text version, and need for cognition. This analysis revealed a decrease in information recalled from the immediate test to the delayed test. There was

an interaction of testing time and text version which showed that the original text was recalled better at the immediate testing and there was no difference between the two versions at the delayed test. Independent of the time of testing, readers with a high need for cognition recalled more information. The second analysis was on the short answer inference questions. A similar effect of testing time was observed here as well. Immediate recall contained more information than the delayed recall. Text version was also significant such that readers of the revised text correctly answered a higher number of inference questions. An interaction between testing time and text version qualified this outcome such that the superiority of the revised version persisted only in the immediate test but disappeared in the delayed test. It was also interesting to note that the readers answered as many questions in the delayed test as they did in the immediate while readers of the revised test answered less in the delayed even though they had answered the same set of questions before. The third and fourth analyses were conducted only on the delayed tests to observe the effects of being tested. In the delayed free recall, there was a superiority of the readers with high need for cognition. In both the delayed free recall and the delayed short answer tests there were complex interactions of text version, need for cognition, and whether the readers were tested before or not. We concluded that the delayed testing provides a different aspect of learning from text revealing that many findings based on immediate testing may be invalidated by findings of the delayed testing. In addition, need for cognition and text structure appear to be influential even though only under certain circumstances.

The Study of Locative Sentences from Corpora of Narratives

M. de Vega*, M. J. Rodrigo, M. Ato, D. Dehn*** & B. Barquero***

***University of La Laguna (Spain)**

****University of Murcia (Spain)**

*****University of Saarbrucken (Germany)**

mdevega@ull.es

This research focuses on the analysis of directional expressions taken from corpora of narratives in two different languages: Spanish and German. These expressions use direction locatives equivalent to the English above/over, below /under, front and behind. For instance:

(1) Spanish examples: La pobre pasó ante las bandejas de pollo asado [the poor (girl) passed in front of the grilled chicken tray] Me oculté tras los troncos [I hide myself under the logs]

(2) German examples: Sie sah die Sterne über den Bäumen [She saw the stars above the trees] Er stand vor dem Schaufenster [He stood in front of the shopping window]

The comprehension of locative expressions posits an important problem of interface between language (a linear code governed by grammar) and spatial representations (a parallel code governed by geometric and dynamic principles). A popular approach among linguists is the hypothesis of schematization (Talmy, 1983; Herskovits, 1985; Landau & Jackendoff, 1993). These authors emphasize the abstract and schematic character of spatial representations built from locative expressions. First, locative prepositions express a gross topological categorization of

space, rather than a fine-grained Euclidean analysis of space. Secondly, the use of locative preposition is independent of any particular geometric feature of the target and the frame. For instance, there are not different locatives depending of the shape, size, mobility, or substance of the target or the frame. Thus, locatives seems to convey and ideal meaning, and the target and the frame are considered as minimal geometric entities. For instance, the target (or figure) of directional expressions can be considered as a point, and the frame (or ground) as a system of coordinate axes (Talmy, 1985). [The only reported exception to this principle in English are the prepositions along and across that requires its frame object to be elongated (see Landau & Jackendoff, 1993)]. A more psychological version of the schematization hypothesis considers that understanding locative expressions relies on the "where" system (the topological relations among objects), whereas information from the "what" system (e.g., sensory-motor features and geometric features) is irrelevant (e.g., Landau & Jackendoff, 1993). The modularity of these systems allows that expressions involving locative prepositions activate selectively the "where" system, independently of the particular features of the target and frame objects.

The integration hypothesis we propose here, postulates that rather than an independence between the use of locative prepositions and the features of objects, we may expect an integration between the "where" and "what" information. The asymmetry generally found between targets and frames of locative expressions may challenge the schematization principle. For instance, "The cat is in front of the house" is acceptable, whereas "The house is behind the cat", is not. Thus, the choice of prepositions seems to be sensible, after all, to some geometric, dynamic or functional properties of the target and the frame (e.g., the frame probably tends to be more static and bigger than the target). On the other hand, it may be possible that the properties of targets and frames also differ systematically depending on the particular direction or dimension tagged by the locative preposition. In summary, this study explores: a) how targets and frames differ in their geometrical, dynamical and functional features; b) what features of targets and/or frames are associated to particular directions (e.g., above, front) or dimensions (e.g., front-back, above-below) expressed by the corresponding locative prepositions. The features encoded both for the target and the frame were: Animacy; Partitiveness; Countability; Solidity; Mobility. In addition, some layout features were encoded: Target-frame surface contact; Target smaller than frame; and Projected point of view (e.g., the car is in behind the mountain). All these binary features were encoded for each locative sentence.

Method

Corpora. We used two corpora of contemporary novels: the Spanish corpus (1 million words) was created by Alameda and Cuetos (1995), and the German corpus of (1 million words) was obtained from the Institut für Deutsche Sprache.

Sampling procedure. Only relativistic expressions including axial terms and two name phrases as target and frame, respectively, were selected. Sentences that had a metaphorical or a non-spatial meaning were discarded. The axial terms in Spanish and German and their translation into English are shown below:

ENGLISH	SPANISH	GERMAN TRANSLATION
Above	sobre, encima de	über
Below	bajo, debajo de	unter
In front of	ante, frente a	vor
Behind	tras, detrás de	hinter

Results

The data were submitted to several logistic regression analyses, using all or some of the features as independent variables, and dimensions (front-back versus right-left), or directions (e.g., front versus back), or object role (target versus frame) as binary dependent variables. Most of the results are convergent in the two languages, in particular:

(1) VERTICAL (above-below) is statistically associated to partitive targets, target-frame contact, and target smaller than frame.

(2) Instead, HORIZONTAL (front-behind) is statistically associated to projected view, animate and solid targets, animate, solid and countable frames.

(3) Some distinctive features of directions were also shared in Spanish and German. For instance, the FRONT involved projected view, countable frames, solid targets and solid frames, whereas ABOVE is more frequently associated to target-frame contact and partitive targets in the two languages.

(4) Concerning the contrast between TARGETS and FRAMES across directions, targets were more frequently animate, and mobile and frames were more frequently solid in Spanish and German.

Discussion

The corpora study demonstrated that authors who write narratives in two different languages are quite systematic in their choice of targets and frames for their locative expressions. Rather than processing independently the semantics of spatial prepositions and the semantic of targets and frames, authors seem to consider the features of targets and frames which are appropriate to the particular dimension or direction tagged by the preposition.

A second study was run to test the integration hypothesis. We designed sentences involving a given preposition in such a way that the target and the frame had some of the typical features that we had discovered for the corresponding dimension. Afterwards, we gave two groups of people those sentences under two different task requirements: 1) A preposition choice task, involving pairs of prepositions from different dimension (e.g., The computer keyboard was BELOW / BEHIND the bar awning); 2) A preposition choice task, involving pairs of prepositions from the same dimension (e.g., The computer keyboard was BELOW / ABOVE the bar awning). The results showed that in the different dimension task participants chose the "correct" preposition most of the times (90%), because they identified the target and frame features as related to the dimension to which the correct preposition belongs. Instead, participants in the same dimension task chose the "correct" preposition almost randomly (54%), because some features of the target and the frame are equally appropriate for the two prepositions which belong to the same dimension.

The results are discussed in terms of the integration hypothesis. In addition, embodiment theories of text meaning (e.g., Glenberg, 1997) and the constructionist and functionalist approaches to grammar (e.g., Goldberg, 1995) are considered.

Processing Figurative Language in the Underspecification Model

S. Frisson* & M. Pickering**

*HCRC Glasgow and University of Antwerp-Ufsia (Scotland)

**University of Edinburgh (Scotland)

steven.frisson@ufsia.ac.be

All the psycholinguistic models that have been proposed in the figurative language literature are concerned with the time-course of sense interpretation. Essential to a correct understanding of these models is their assumption of what sense is activated *immediately* upon encountering a word that can or needs to be interpreted figuratively: Is a figurative interpretation only possible via the literal sense (*literal-first*), can the figurative sense be activated without the literal sense (*direct access or figurative-first*), is the most salient sense activated first, or are all senses activated at the same time (*parallel*)? Since the details of these models differ precisely at the earliest stages of processing, it is crucial to investigate the on-line processing of sense interpretation (see also Dascal, 1989; Gibbs, 1992). Surprisingly, however, few data are available.

In this paper, we discuss three eye-tracking experiments investigating the on-line semantic interpretation of sentences containing words with multiple (semantically related) *senses*. To account for the data, a general model is proposed (the Underspecification Model) which holds that, *initially*, one underspecified meaning rather than a specific sense is activated for these kinds of words. Only in a later stage do readers "home-in" on the contextually intended sense. A fourth eye-tracking experiment concerned the processing of verbs that have multiple (semantically unrelated) meanings.

Experiments 1 and 2 (see Frisson & Pickering, 1999) deal with the processing of *established* metonymies. The eye movement data show that early processing of place-for-institution metonymies (e.g., *The blasphemous woman had to answer to the convent*) does not differ from their literal counterparts (e.g., *These two businessmen tried to purchase the convent*), though processing is disrupted when the edifice has no established metonymic sense (*e.g., answer to the stadium*). Aside from *literalness*, the relative frequency of the two senses and the "basicness" of one sense over the other also did not affect early processing times. A comparable pattern was found for place-for-event metonymies, with no initial processing differences between metonymic expressions like *protesting during Vietnam* and literal expressions like *hitchhiking around Vietnam*, even though the literal interpretation was almost three times as frequent as the metonymic interpretation. Models of figurative language processing which hold that processing occurs via the activation of *one specific* sense (e.g., the literal, the most frequent, or the most basic sense) cannot explain these data. In contrast, we argue that readers initially activate an abstract, underspecified meaning of these words, which encompasses both senses (as well as all other semantically related senses that are known by them) and then uses contextual information to specify the interpretation.

Experiment 3 (see Pickering & Frisson, in press) deals with the processing of verbs with multiple senses, one of which being a conventionalised metaphorical sense. The processing of metaphorical expressions like *He disarmed almost every critic* was compared to the processing of literal expressions like *We disarmed almost every rebel*. These expressions were either preceded by a context supporting one interpretation or a semantically neutral context. The eye movement data did not show any immediate beneficial effect of supportive preceding contexts. In addition, and in agreement with the findings for the metonymic expressions, no immediate differences were found between the literal and the metaphorical interpretation, even though the metaphorical interpretation was considerably lower in frequency. However, significant differences were found for later processing measures, indicating that homing-in on a figurative sense of a verb requires more effort than when that verb expresses the literal sense. Although it is unclear whether this later effect is caused by frequency differences between the two senses, differences in saliency, or a general bias towards a literal interpretation over a metaphorical, the crucial finding is the *lack* of any *immediate* effects, even when the preceding context is semantically neutral. Again, we argue that the Underspecification Model can best capture these results.

In a fourth eye-tracking experiment, homonymous verbs with multiple (semantically unrelated) *meanings* were tested. For these verbs, no *single* (informative) underspecified meaning can be set up since the overlap between the different interpretations is too small. In a design similar to the previous experiment, homonymous verbs like rule (*a country* vs. *a line*) were either preceded by a supportive or a neutral context. In contrast to the findings for the verbs with multiple senses, supportive contexts *did* show an early beneficial effect, probably in order to help selecting between the two incompatible meanings. Moreover, we observed a (slight) interaction around the noun region, indicating that the subordinate interpretation preceded by a neutral context was harder to process than the other three conditions (a finding that was absent for the verbs with multiple senses). In addition, when direct comparisons were performed on the data of both experiments, it appeared that verbs with multiple senses were easier to process than verbs with multiple meanings. This supports the notion that verbs of the former type activate only one (underspecified) meaning while homonymous verbs activate all (underspecified) meanings and a time-consuming selection needs to be made between them.

Consistent with our own studies, results of other on-line studies (Frazier & Rayner, 1990; Frazier, Pacht, & Rayner, 1999; McElree & Nordlie, 1999) can be interpreted in a model that posits the initial activation of underspecified meanings. We argue that contrasting results, stemming from less sensitive techniques, indicating the immediate activation of a specified sense, are unreliable since no distinction is made between early and late interpretation processes. Finally, we argue that the Underspecification Model does not rely (explicitly or implicitly) on a functional difference between literal and figurative language. Rather, figurative senses do not seem to behave differently from any other semantically related sense of a word. Moreover, the hypothesis that we initially activate a single, underspecified meaning questions the representation of separate senses in the mental lexicon, and provides a more psycholinguistically plausible account of incremental semantic processing.

References

Dascal, M. (1987). Defending literal meaning. Cognitive Science, 11, 259-281.

Frazier, L., Pacht, J. M., & Rayner, K. (1999). Taking on semantic commitments II: Collective versus distributive readings. Cognition, 70, 87-104.

Frazier, L., & Rayner, K. (1990). Taking on semantic commitments: processing multiple meanings vs. multiple senses. Journal of Memory and Language, 29, 181-200.

Frisson, S., & Pickering, M. J. (1999). The processing of metonymy: Evidence from eye movements. Journal of Experimental Psychology: Learning, Memory, and Cognition, 25, 1347-1365.

Gibbs, R. W. (1992). When is metaphor? The idea of understanding in theories of metaphor. Poetics Today, 13, 575-606.

McElree, B., & Nordlie, J. (1999). Literal and figurative interpretations are computed in equal time. Psychonomic Bulletin & Review, 6, 486-494.

Pickering, M. J., & Frisson, S. (in press). Processing ambiguous verbs: Evidence from eye movements. Journal of Experimental Psychology: Learning, Memory, and Cognition.

Metaphor and the Space Structuring Model: Evidence from Event-Related Brain Potentials

S. Coulson

University of California, San Diego (U.S.A.)

coulson@cogsci.ucsd.edu

In this talk we outline the meaning construction operations involved in metaphor comprehension, and assess the claim that the right hemisphere (RH) is specialized for this sort of nonliteral processing. The focus is on contrasting predictions for the on-line comprehension of metaphors made by two models of high-level language processing. One model is the standard pragmatic model (Grice, 1975), which posits distinct mechanisms for literal and nonliteral language processing. The other model is the space structuring model, and is based on the theory of conceptual integration, also known as blending (Coulson, in press; Fauconnier & Turner, 1998). In the space structuring model, both literal and nonliteral comprehension proceed via the construction of simple cognitive models and the establishment of various sorts of mappings, or systematic correspondences between elements and relations in each.

Standard Pragmatic Model. Premised on a firm distinction between semantics and pragmatics, the standard pragmatic model holds that literal language is processed by compositional mechanisms, while metaphors require pragmatic inference and the appeal to background knowledge. Besides the intuition that metaphor comprehension involves different sorts of knowledge than the comprehension of more literal language, the chief source of support for this view comes from neuropsychology, and the observation that patients with damage to the right cerebral hemisphere (RH) present with high-level language deficits, including difficulty understanding metaphors. This contrasts with the prevalent view among psycholinguists that the processing of literal and nonliteral language involves qualitatively similar sorts of operations.

Space Structuring Model. In the space structuring model, all language comprehension involves the construction of multiple cognitive models and the establishment of mappings between their components based on identity, similarity, and analogy (Coulson, in press; Fauconnier & Turner, 1998). Consequently, metaphoricity should be viewed as a continuum, with literal meanings and lexicalized metaphors at one end, and completely novel metaphors at the other. Moreover, this model suggests a middle ground between the literal and the metaphoric that involves the same

sort of mapping and blending operations as in metaphor. For example, the word "gem" could be used in the following sentences:

(a) After, giving it some thought, I realized the new idea is a gem.

(b) That stone we saw in the natural history museum is a gem.

(c) The ring was made of tin, with a pebble instead of a gem.

In (a), a metaphoric use, the speaker's idea is mapped analogically onto the concept of a gem. In (b), the literal use, a stone seen in a museum is subsumed under the category *gem* via an identity mapping. In (c), a use we refer to as " literal mapping ", the pebble in the tin ring has been mapped onto the gem in a more expensive ring. While the term "gem" is used in (c) to convey its conventional literal meaning, the mapping exploits a correspondence of abstract relational structure between the concept of a prototypical ring and the cheaper, pebble ring described in the sentence. Conceptual integration theory suggests that while (b) and (c) both employ literal uses of the same word, the blending operations in (c) are more complex. Consequently, gem" in (c) should be slightly harder to comprehend than the literal use in (b), and slightly easier to comprehend than the metaphoric use in (a).

Our experiments addressed three issues: (i) Whether there is a qualitative difference in the processing of metaphors and more literal language; (ii) whether the continuum of metaphoricity described above predicted on-line comprehension difficulty; and, (iii) whether the right hemisphere is specialized for metaphor processing. We utilized a direct measure of brain activity that occurs in on-line comprehension: Event-related brain potentials (ERPs). Besides assessing whether literal and metaphoric stimuli elicited qualitatively different ERPs, we capitalized on established properties of an ERP component, the N400, whose amplitude is known to be inversely proportional to the difficulty of semantic integration. By measuring the amplitude of this component, we were able to assess the difficulty of comprehending literal and figurative language in the various conditions.

Experiment 1. In the first experiment, ERPs were recorded from normal adults as they read sentences like those in examples (a)-(c) above. ERPs, measured from 300-500 ms after the onset of the sentence-final words, differed quantitatively, but not qualitatively, as a function of metaphoricity. Literal endings, as in example (b), elicited the least N400, suggesting they were the easiest to comprehend. Literal mappings, as in example (c), elicited slightly larger N400s. Metaphoric endings, like (a), elicited the largest N400 components, suggesting they were the hardest to understand. This graded pattern of N400 amplitudes is consistent with conceptual integration theory which posits a continuum from conventional meanings to novel metaphors.

Experiment 2. Further, in contrast to the patient data which suggests the RH plays an important role in the interpretation of metaphors, results from our first experiment did not suggest differential involvement of the RH in the processing of metaphoric language. However, to more closely examine interhemispheric differences in on-line metaphor processing, we recorded ERPs as participants read sentences that ended with words presented in the left or right visual field. Lateralized presentation was done to selectively stimulate each cerebral hemisphere. Though information presented in the center of gaze is projected to both sides of the brain, extra-foveal information is projected only to the contralateral hemisphere. Consequently, words presented to the left visual field would be processed first by structures in the RH, while words presented to the right visual field would be processed first by structures in the LH.

Stimuli included three sorts of sentences: *hicloze*, that is, sentences which ended predictably, *literal* sentences which ended with an unpredictable word used in its literal sense, and *metaphoric* sentences which ended with an equally low cloze word used metaphorically. If the RH is specialized for metaphor processing, we should expect metaphor effects (the difference in

N400 amplitude elicited by literal versus metaphoric uses of the same word) to be more pronounced with presentation to the right visual field (LH) than to the left visual field (RH).

In both visual fields, hicloze endings elicited smaller N400s than either of the less predictable ending types (literal or metaphoric). When presented to the right visual field (LH), the last words of both literal and metaphoric sentences elicited fairly similar N400s. In contrast, left visual field (RH) presentation resulted in larger N400s for metaphoric than literal sentences. Results suggest interhemispheric differences are indeed relevant to the semantic integration of metaphoric words, but argue against the simple hypothesis that the right hemisphere has an advantage processing metaphoric meanings. Larger N400 metaphoricity effects with presentation to the left visual field suggest the RH encounters more rather than less difficulty with metaphoric stimuli.

References

Coulson, S. (in press). Semantic leaps : Frame-shifting and conceptual blending in meaning construction. Cambridge and New York: Cambridge University Press.

Fauconnier, G. & Turner, M. (1998). Conceptual Integration Networks. Cognitive Science, 22, 133-187.

Grice, H. P. (1975). Logic and conversation. In P. Cole and J.L. Morgan (Eds.), Syntax and semantics, Vol. 3, Speech acts. New York: Academic Press.

Figurative Language: Retention vs. Suppression

R. Giora

Tel Aviv University (Israel)

giorar@post.tau.ac.il

According to the graded salience hypothesis (Giora, 1997, 1999, in press), salient meanings (meanings coded in the mental lexicon, enjoying familiarity, conventionality, frequency, and prototypicality) would be activated initially, regardless of contextual information (Giora & Fein, 1999b; Giora, Fein & Schwartz, 1998; Swinney, 1979; Rayner et al., 1994; Peleg et al. in prep.). Thus, comprehending *Body & Sole* - the name of a shoe store - involves activating the salient, "soul" sense of the homophone, regardless of its contextual inappropriateness. Would such meanings be discarded as "irrelevant" as might be deduced from studies showing suppression of contextually incompatible meanings (cf. Gernsbacher, 1990; Swinney, 1979)?

According to the retention hypothesis (Giora, in press; Giora & Fein, 1999a), contextually incompatible meanings would not be discarded in case they have a role in constructing the intended meaning. They would, if they interfere with comprehension. Thus, the "soul" sense induced by *Body & Sole* would be retained to allow for a comparison between *sole* and 'soul', so that the implicature (that the goods is as essential to one's body as the soul is) be derived.

In a number of studies we tested the retention hypothesis regarding irony (Giora & Fein, 1999b; Giora, Fein & Schwartz, 1998) and joke (Colston, Giora & Katz, in prep.) comprehension. We showed that salient, albeit contextually incompatible meanings that are functional in constructing

the intended meaning are not suppressed upon the derivation of the intended meaning. Irrelevant meanings that are not conducive to comprehension undergo suppression.

Irony comprehension

According to the graded salience hypothesis the contextually incompatible literal meaning of irony is processed initially on account of its high salience (compared to the low salience of the ironic interpretation). According to the indirect negation view (Giora, 1995), the joint pretense (Clark, 1996; Clark and Gerrig, 1984), and the tinge hypotheses (Dews and Winner, 1999), this contextually incompatible literal meaning should be retained, because it is supportive of and instrumental in constructing the ironic meaning. In contrast, the "standard pragmatic model" (Grice, 1975; Searle, 1979) assumes suppression of this meaning. While it requires that the contextually incompatible literal meaning be processed initially on account of its literality, it further predicts that this inappropriate meaning would be rejected and replaced with the appropriate ironic interpretation.

One way to test the retention versus suppression hypotheses is to examine the temporal aspects of irony comprehension. In Giora and Fein (1999b) we measured response times to literally and ironically related test words after two interstimulus intervals (ISIs). While a short delay (150 msec) should tap activation, a longer delay (1000 msec) might be revealing as to whether the activated meanings have been suppressed or retained. Our findings demonstrate that both salient (familiar) and nonsalient (unfamiliar) ironies are accessed literally initially. They further demonstrate that, consistently with the retention hypothesis, but contra the standard pragmatic model, the literal meaning of irony was not suppressed when the ironic meaning emerged, but remained active even after a long delay (1000 msec). In Giora et al., (1998), we tested the retention hypothesis even under a longer delay condition. But even after an interstimulus interval of 2000 msec, the literal meaning of the utterance embedded in the ironically biasing context was still as active. In contrast, after such a delay, the literal meaning of the same utterance embedded in the literally biasing context began to fade. Once comprehended, it was no longer required for further processes.

Joke comprehension

Jokes are a form of discourse that best exposes our tendency to activate the salient interpretation first. For instance, when Monty Python talked about *baby snatchers*(*And Now to Something Completely Different*), they manipulated our tendency to activate the more salient interpretation initially in order to catch us by surprise when eventually we find out it was the babies who did the snatching (of adults). Similarly, we are caught entirely off guard when we realize that a person who has a *drinking problem* is, in fact, one who spills the liquid all over himself while drinking (*Flying High* by Abrahams, Zucker & Zucker). Jokes, then, involve some salience imbalance which entices the comprehender to process the more salient but eventually unintended meaning first ("snatching babies", '"drinking alcohol excessively") in order to dispense with it and activate a less salient but congruent meaning. To lead our salience-bound mind down the "garden path", the initial context of a joke is usually unambiguous, enhancing the salient (rather than the intended) meaning, so that this interpretation is retained up until the punchline, at which point a sudden incongruity forces reinterpretation (cf. 1 taken from Coulson & Kutas, 1998):

(1)   By the time Mary had her fourteenth child, she'd finally run out of names to call her husband.

According to the retention hypothesis, then, understanding jokes (that do not hinge on punning) should differ from irony comprehension: It should involve suppression of irrelevant meanings, since these meaning are not conducive to the joke interpretation (see also Gernsbacher & Robertson, 1995;

Giora, 1991). In Colston, Giora and Katz (in prep.) we tested this hypothesis. We presented participants with 35 (contrived) jokes (e.g., 2 below) and 70 nonjoke fillers. The jokes involved an ambiguous word (*boxer*)whose less salient (DOG) meaning is the contextually appropriate meaning at the punchline position:

(2)   Muhammad was a great boxer*

        he was never known to bark*

        or wet the floor.*

Probes (displayed at *): Salient meaning: FIGHTER; Less salient meaning: DOG; Unrelated: STOVE; Nonword: VOCK

Participants self-paced their reading of jokes (and nonjokes) which were displayed line by line on a computer monitor. They had to make lexical decisions as to whether a probe, displayed on the screen immediately (0 msec. ISI) after offset of each line, was a word or a nonword in English. Results indeed demonstrate that, as anticipated, the (contextually compatible) salient meaning (FIGHTER) of the critical word (*boxer*) was available at the first position. However, as the text proceeded, this meaning was deactivated and was no longer available at the punchline position. These findings demonstrate that contextually incompatible meanings are involved in comprehension on account of their saliency. Still, they are not discarded automatically. They are retained if they are instrumental in constructing the intended meaning (e.g., in constructing the ironic meaning of unfamiliar ironies). They are suppressed if they are not (as in the case of jokes).

Convention, Form, and Figurative Language Processing

B. F. Bowdle* & D. Gentner**

*Indiana University (U.S.A.)

**Northwestern University (U.S.A.)

bbowdle@indiana.edu

Metaphors serve to create mappings between concepts from disparate domains of knowledge. For example, in the metaphor *The mind is a computer*, an abstract entity is described in terms of a complex electronic device. But how are such mappings established during metaphor comprehension? At present, there are two prominent competing approaches to this question. The first approach treats metaphors as comparison statements, in which the target (first term) and base (second term) are understood at roughly the same level of abstraction, and matches between the two concepts are made directly (e.g., Miller, 1979; Ortony, 1979; Tversky, 1977). The second approach treats metaphors as categorization statements, in which the target concept is understood as a member of a superordinate category generated from the base concept (e.g., Glucksberg & Keysar, 1990; Glucksberg, McGlone, & Manfredi, 1997; Honeck, Kibler, & Firment, 1987; Kennedy, 1990). In the first part of this presentation, we briefly review these two views of metaphor comprehension, pointing out strengths and limitations of each. We then propose a hybrid approach

– *the career of metaphor hypothesis* – according to which different types of metaphoric mappings are established at different levels of conventionality (Bowdle, 1998; Bowdle & Gentner, 1995, 1999; Gentner & Wolff, 1997). On this view, novel metaphors are processed as analogical comparisons, in which (1) the target and base representations are structurally aligned, and (2) further predicates from the base that are connected to the common system are projected to the target as candidate inferences. Over time, however, such mappings can lead to the creation of stable metaphoric categories. Specifically, if the same base concept is repeatedly aligned with different targets so as to yield the same basic interpretation, then the common system may become conventionally associated with the base as an abstract schema. At this point, the base term will be polysemous, having both a domain-specific meaning and a related domain-general meaning. As metaphors become conventionalized, then, there should be a shift in mode of mapping: While novel metaphors are processed as comparisons, by aligning the target concept with the literal base concept, conventional metaphors may be processed as categorizations, by aligning the target concept with the abstract schema named by the base term.

In the second part of this presentation, we describe some recent studies from our laboratories investigating the career of metaphor hypothesis. Central to the logic of these studies is the distinction between metaphors and similes. As is well known, nominal metaphors (figurative statements of the form *X is Y*) can typically be paraphrased as similes (figurative statements of the form *X is like Y*). For example, one can say both *The mind is a computer* and *The mind is like a computer*. While these two forms are often viewed as synonymous, this linguistic alternation offers a potentially valuable tool for examining the use of comparison and categorization during figurative language comprehension. This is because metaphors are grammatically identical to literal categorization statements (e.g., *A sparrow is a bird*), and similes are grammatically identical to literal comparison statements (e.g., *A sparrow is like a robin*). Thus, assuming that form reflects function in both literal and figurative language, metaphors and similes should tend to promote different comprehension processes. Specifically, metaphors should invite classifying the target as a member of a category named by the base, whereas similes should invite comparing the target to the base.

Across a wide range of experiments involving the metaphor-simile distinction, we found strong support for the claim that conventionalization results in a shift from comparison processing to categorization processing. In production tasks, we found that subjects preferred expressing novel figurative statements as similes, but that there was a shift in preference towards the metaphor form as the figurative statements became increasingly conventional. In on-line comprehension tasks, we found that similes were interpreted faster than metaphors for novel statements, but that metaphors were interpreted faster than similes for conventional statements. Finally, in recall tasks, we found that metaphors were more likely to be misremembered as similes than the reverse when the statements were novel, but that similes were more likely to be misremembered as metaphors than the reverse when the statements were conventional.

In the final part of this presentation, we relate the career of metaphor hypothesis to other theoretical approaches to the role of conventionality in figurative language comprehension (e.g., Blank, 1988; Blasko & Connine, 1993; Giora, 1997). We also discuss how conventionality and grammatical form may jointly determine (1) whether figurative statements are processed directly or indirectly, and (2) whether metaphoric mappings are established between individual concepts or entire conceptual domains.

Cognitive Evidence for a Parameterization of Cohesion and Coherence relations

M. Louwerse

University of Edinburgh (Scotland)

max@ling.ed.ac.uk

Most theories of text comprehension assume that comprehenders build a coherent mental representation of events whose relationships are cued bycohesive linguistic devices in a text. I ask two research questions: (1) What is the effect of the various cohesion categories on the comprehension process? (2) Where does the construction of coherence take place in the comprehension process?

It is proposed that the answers to these questions depend on the TYPE and POLARITY of cohesive expressions (see Sanders, Spooren & Noordman, 1992, 1993 and Knott, 1996). TYPE contains the categories CAUSAL, TEMPORAL and ADDITIVE; POLARITY contains the categories POSITIVE and NEGATIVE. POSITIVE relations extend, while NEGATIVE relations cease to extend the situation expressed by the preceding/succeeding clause. TYPE relations refer to time and causality (CAUSAL); time only (TEMPORAL), or neither (ADDITIVE). The factorial combination of TYPE and POLARITY offers six kinds of connective relations, realized in common English connectives.

	CAUSAL	TEMPORAL	ADDITIVE
POSITIVE	because	before	moreover
NEGATIVE	although	until	however

Several studies have investigated how textual cohesion supports representational coherence for both POSITIVE and NEGATIVE relations. CAUSAL relations are processed faster than TEMPORAL which are faster than ADDITIVE relations (Caron et al., 1988; Deaton & Gernsbacher, in press; Sanders & Noordman, 2000). POSITIVE relations are generally processed faster than NEGATIVE relations (Townsend, 1983). The problem with many of these studies is that the materials are biased towards a particular type of coherence. To answer the first research question with regard to the effects of cohesion categories on the comprehension process, we need texts that do not bias the comprehender in establishing a coherence relation.

In addition to differences in processing speed, Townsend and Bever (1978) and Townsend (1983) argue that there are also differences in the processes that the various relations induce. They showed that questions using different cohesion cues directed the subjects' attention to either structural or associative information. Causal explicitness increased accessibility to associative information. Also, it increased the reaction times of late targets and decreased those of early targets.

The finding that different cohesion relations direct resources to different processes at different stages in the comprehension process, brings us back to the second research question where processing takes place. Millis and Just (1994) argue that the comprehender postpones the construction of the mental representation until the final (wrapping-up) stage of the comprehension process. Traxler et al., (1997) on the other hand argue that coherence relations are built incrementally throughout the comprehension process.

I propose that the TYPE and POLARITY relations determine not only processing speed (research question 1), but also where that processing will occur (research question 2). To answer these two questions I first prepared unbiased materials. In a data-collection experiment I used texts like the

following: " My neighbor played saxophone. I didn't like it [conjunction] he practiced every day ". The second and third clause can be conjoined by any of the conjunctions 'because', 'although', 'later', 'until', 'moreover' and 'however', each of which occupies a different cell of the type x polarity classification. Informants were asked to judge how natural each text was with one connective and all items were seen with all connectives by an equal number of judges.

A selection of unbiased texts was made to be presented in a reading time experiment with each connective. Because the same texts were used throughout, unwanted contextual effects of frequency, word length, syntax, morphology, are unlikely.

In comparing reading times within subjects and texts, I predicted TYPE and POLARITY effects: that the two specified relations, CAUSAL and TEMPORAL, would be processed faster than ADDITIVE relations, for which some CAUSAL or TEMPORAL relationship might have to be inferred by the reader. Furthermore, I followed the literature in predicting that POSITIVE relations are processed faster than NEGATIVE relations. I also predicted a TYPE effect on the point in reading where additional time is needed. Because ADDITIVE relations demand inference about cause or time, they should induce an incremental comprehension process, because comprehenders need information from the text to specify the coherence relation to establish. More specified cohesion relations, however, could be processed towards the end of the comprehension process, since the desired form of coherence is already specified by the connective. Accordingly, I predicted that CAUSAL relations, being more specific than ADDITIVE relations, are processed faster towards the beginning but not toward the end of the clause, with TEMPORAL relations in between. Similarly, I predicted that POSITIVE relations are processed faster than NEGATIVE relations toward the beginning, but not toward the end of the clause.

Some 36 texts, which subjects judged natural in all six cohesion conditions, were presented for non-cumulative self-paced reading with 18 fillers. Seventy-two subjects were assigned to text-connective combinations by Latin Square. They read the experimental texts and answered a comprehension question on each. Before analysis, wrong answers to questions, which were equally distributed over all conditions, texts and subjects, were discarded. Next outliers beyond 2.5 SD from the mean were replaced by the mean for each condition. As expected, no effects were found for either TYPE or POLARITY on the reading times for the first or second clause (all Fs < 1). Contrary to what was predicted, reading times of the third clause showed no significant effect either. Analyses were then run for the first, second, third and final word of the third (post-connective) clause, as only these could be found in all texts. To account for differences in length and frequency of the conjunctions, an ANOVA was carried out on the residual reading times of the conjunctions. An effect was found for TYPE ($F1(2, 142) = 4.72$, $p = .010$; $F2(2, 62) = 4.38$, $p = .017$), but not for POLARITY. As predicted, CAUSAL relations were processed faster than TEMPORAL relations, which were processed faster than ADDITIVE relations. The word after the conjunction showed an identical pattern, with an effect only for TYPE ($F1(2, 142) = 7.01$, $p = .001$; $F2(2, 62) = 5.088$, $p = .009$). The initial strong effects for TYPE disappeared in the second word after the conjunction, with only a by-subjects effect for TYPE ($F1(2, 142) = 3.43$, $p = .035$). The final word of the third clause again showed an effect for TYPE ($F1(2, 142) = 4.45$, $p = .013$; $F2(2, 62) = 3.38$, $p = .04$). This time, however, the pattern had reversed. CAUSAL (fully specified) relations were processed slowest here, then TEMPORAL, then ADDITVE. This delayed coherence was what was predicted for those relations that are specific. No effects were found for POLARITY.

Like the results found in the literature, our experiment shows overall processing speed CAUSAL-TEMPORAL-ADDITIVE. But this depends on the position in the comprehension process. ADDITIVE relations are processed incrementally, with a processing load at the beginning of the clause. CAUSAL relations are processed in a wrapping-up stage toward the end of the clause. No effects were found for POLARITY. This experiment provides cognitive evidence for the proposed types of coherence relations and for a set of processes based on types.

Coherence and Evidence in Testimony Evaluation on Incest Narratives

B. Klettke & A. C. Graesser

The University of Memphis (U.S.A.)

bklettke@memphis.edu

Narratives play an important role in the legal domain. Cases that make it to court tend to be those that contain controversial evidence about a series of events that occurred in the past. More specifically, child sexual abuse cases are often based on little physical evidence, and may be solely based on the child's claims (Mason, 1991; Kovera & Bordiga, 1996). Research on jury decision making has shown that jurors create a story (narrative) to make sense of the information, and that both evidence quality and story coherence have a major impact on their representation of the case (Pennington & Hastie, 1986, 1988, 1990, 1993a, 1993b). Available studies have supported the importance of narrative construction, but the status of evidence quality is not clearcut. Pennington and Hastie claim that the construction of the story is primarily directed by the explicit evidence (in conjunction with world knowledge and story structure knowledge), whereas coherence (comprised of consistency and completeness) determines the acceptability of the story. Other research (Kuhn, Weinstock, & Flaton, 1994; Brem & Rips, 1996; Voss, Wiley, & Sandack, 1999), however, has shown that evidence per se might play a lesser role and may be overridden by coherence in jury decision making. Voss et al., (1999) reported that poor narrative coherence by the prosecution leads to fewer guilty verdicts.

The present study examined the extent to which story coherence and the quality of the evidence have an effect on the interpretation of testimony on sexual abuse. Sixty three college students rated four different testimonies in the form of narratives. The narratives were extracted from a popular book called " Voices in the Night: Women Speaking About Incest " (McNaron & Morgan, 1982). The independent variables (evidence) and (coherence) were manipulated into two conditions (high versus low), yielding four possible conditions: High evidence / high coherence, high evidence / low coherence, low evidence / high coherence, low evidence / low coherence. Four critical sentences in each narrative were manipulated (2 for evidence and 2 for coherence) whereas most of the story was the held constant across conditions. High versus low coherence was defined by the absence versus presence of contradictory information. High quality evidence was based on physical evidence or objective information whereas low quality evidence reflected mere opinion expressed by one of the characters in the narrative. The assignment of condition to stories was counterbalanced across subjects.

After reading the stories, the subjects filled out questionnaires with three questions pertaining to the dependent variables: Guilt, contradiction, and plausibility. The first question measured the perceived guilt of the offender on a six-point scale (6 = guilty, 1 = innocent). The second question assessed the perceived amount of contradiction in the story on a six-point scale (6 = non-contradictory, 1 = contradictory). The third question assessed the overall plausibility of the testimony on a six-point scale (6 = plausible, 1=implausible). Three subjects were dropped from the analysis due to missing data.

An analysis of variance performed on the dependent variable guilt, yielded a significant main effect for coherence, $F(1, 59) = 3.26$, $p < .05$, one-tailed; neither evidence, nor the interaction were significant. For the dependent variable contradiction, an ANOVA showed a statistically significant main effect for coherence, $F(1, 59) = 3.62$, $p < .05$, one-tailed. Evidence was almost significant, $F(1, 59) = 2.74$, $p = .052$, one-tailed, whereas the interaction was not significant. For the dependent variable plausibility, an ANOVA revealed a significant main effect for coherence, $F(1, 59) = 4.71$, $p < .05$, one-tailed), with neither evidence nor the interaction being statistically

significant. In these analyses, the high coherent versions were viewed as involving higher guilt, fewer contradictions, and higher plausibility. When the three dependent variables were collapsed to form an overall effect labeled as " impact ", an ANOVA showed a significant main effect for coherence, $F(1, 59) = 6.94$, $p < .05$, but no significant effects of evidence quality and the interaction. These results support the claim that coherence is more important than evidence when evaluating testimonies on sexual abuse. Participants were more strongly persuaded by coherence on all three dependent variables. In other words, the present findings support the hypothesis that coherence may be the overriding factor in testimony evaluation and jury decision making.

In addition, these results indicate that jurors need coherence in order to construct the evidence into a narrative, similar to the theory that Pennington and Hastie (1986, 1988, 1900, 1993a, 1993b) have proposed. Coherence seems to be a crucial factor for narrative quality. In addition, since coherence was manipulated by the presence/absence of a contradiction, coherence seems to be more effective when it is broad and explanatory. This hypothesis is also supported by Read and Marcus-Newhall's (1993) study on explanatory coherence. Results of this study indicated that when a single fact needed to be explained a narrow explanation was rated better. However, when multiple facts needed to be explained, a broad explanation was rated as more appropriate. The results indicated that explanations are often based on causal chains. In a courtroom scenario this means that the lawyer or prosecutor should try to integrate the evidence into a narrative that provides causal and explanatory coherence.

References

Brem, S., & Rips, L. (1996). Explanation and evidence in informal reasoning. In W.G. Cotrell (Ed.), Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society, 271-276. Mahwah, NJ: Lawrence Erlbaum.

Kovera, M.B., & Bordiga, E. (1996). Children on the witness stand: The use of expert testimony and other procedural innovations in U.S. child sexual abuse trials. In B.L. Bottoms and G.S. Goodman, (Eds.),International perspectives on child abuse and children's testimony : Psychological research and law. Thousand Oaks: Sage Publications.

Kuhn, D., Weinstock, M., & Flaton, R. (1994). How well do jurors reason? Competence dimensions of individual variation in a juror reasoning task. Psychological Science, 5, 289-196.

Mason, M.A. (1991). A judicial dilemma: Expert witness testimony in child sexual abuse cases. Journal of Psychiatry and Law, 19, 185-219.

McNaron, T. A., & Morgan, Y. (Eds.) (1982). Voices in the night : Women speaking about incest. Pittsburgh: Cleis Press.

Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision making. Journal of Personality and Social Psychology, 51, 242-258.

Pennington, N., & Hastie, R. (1988).Explanation-based decision making: Effects of memory structure and on judgment. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14, 521-533.

Pennington, N., & Hastie, R. (1990). Practical implications of psychological research on juror and jury decision making. Personality and Social Psychology Bulletin, 16, 90-105.

Pennington, N & Hastie, R. (1993a). The story model for juror decision making. In R. Hastie (Ed.), Inside the juror: The psychology of juror decision making. Cambridge: Cambridge University Press.

Pennington, N., & Hastie, R. (1993b). Reasoning in explanation-based decision making.Cognition, 9, 123-163.

Read, S.J., & Marcus-Newhall, A. (1993).Explanatory coherence in social explanations: A parallel dustributed processing account. Journal of Personality and Social Psychology, 65, 429-447.

Voss, J. F., Wiley, J., & Sandak, R. (1999). On the use of argument. In S.R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), Narrative comprehension, causality, and coherence: Essays in honor of Tom Trabasso. Mahwah, NJ: Erlbaum Publishing Company.

Generating Inferences from Scientific Text

J. Wiley* & J. L. Myers**

*Washington State University, Vancouver (Canada)

**University of Massachusetts, Amherst (U.S.A.)

wiley@vancouver.wsu.edu

While there have been many studies on inference generation from narrative text, there has been comparatively less research on expository or scientific texts. Recently there have been a number of studies on the topic. Unfortunately, no clear consensus emerges from these studies; Noordman, Vonk, and Kempf (1992) failed to find evidence for on-line scientific inferences, whereas Singer and his colleagues (1997, 1999) and Millis and Graesser (1994) found evidence that spontaneous inferences about causal antecedents are made routinely by readers of scientific text.

In the Noordman et al. study, readers were presented with passages that either contained two premises and a conclusion, or just a target sentence containing the conclusion and one premise. They predicted that if readers generate inferences on-line then reading time should take longer on the target sentence when the implicit premise needed to be computed. Similarly, if implicit premises are computed by readers, then verification times on those implicit premises presented AFTER the target sentence should not differ across conditions. However, they found that reading times on the target sentence did not differ between conditions. Further, verification times on the implicit premise were longer in the implicit condition. This suggests that students do not infer the missing premise spontaneously. However, when given specific reading instructions, readers in the implicit condition did take longer on the target sentence, and took no longer to verify the missing premise after reading the passage. These studies suggest that readers do not make causal inferences spontaneously while reading expository text, but that they can be made when the reader's goal requires it. The findings of Millis and Graesser led to a different conclusion. In their study, readers made faster lexical decisions when the probe word was related to a causal inference than when the probe word was unrelated. This suggests that readers generated causal inferences on-line. There are a number of possible explanations for the discrepancies between these studies. The most obvious difference is the nature of the task that was used to assess inference generation. Millis and Graesser employed a probe task. Probe tasks, especially those requiring the reader to make a decision, have been associated with a number of potential confounds and problems. Perhaps the most often-cited problem with probe tasks is that results that look like "priming" can actually result from context-checking, or the compatibility of the probe with the preceding text.

This may be most likely when the reader needs to make a decision about the probe word, and less likely with a task like naming, in which the response is quick enough that the context does not have a chance to effect it. Thus, one interesting question is whether the same "priming" would be observed on naming tasks or on tasks with shorter response times.

Experiment 1

In a series of studies we attempted to replicate the Millis and Graesser findings. Using their materials, we were unable to find an effect of priming using a lexical decision paradigm.

	Wiley & Myers		Millis & Graesser	
	Related	Unrelated	Related	Unrelated
RSVP 500 SOA	654	655	961	921*
RSVP 1000 SOA	650	643	968	928*

One noticeable difference between our results was the average length of our lexical decision times. This suggests that the "priming" effects that were obtained may depend on longer decision times, which would be consistent with a context-checking explanation. The longer response times also opens the possibility that instead of responding more quickly to probes because an inference has already been computed, the probe itself may PROMPT the generation or validation of the inference related to that word (Wiley, Mason & Myers, submitted).

Experiment 2

Given the problem with using probe tasks, a second approach is to look at reading times on sentences related to inferences. In a study based on Noordman et al., Singer (1997, 1999) has found evidence for on-line generation of missing premises on target sentence reading times. However, since both the second premise and the conclusion were part of the target sentence, we do not know whether having both premises allows the reader to generate the inference, thereby speeding the reading of the rest of the sentence, or if having both premises makes it easier for the reader to validate the conclusion once they have read it. Thus in our study, both of the premises and the conclusion are presented as separate sentences. Second, in order to discriminate faster reading due to lack of understanding from faster reading due to better understanding, we chose to use a contradiction paradigm. Readers were presented with either one or both premises, and either an inconsistent or consistent conclusion as a target sentence. Longer reading times on inconsistent conclusions are taken to indicate that the opposite conclusion has been drawn. If readers do not slow down on inconsistent conclusions, it suggests that no inference has been generated. Although previous experiments have found that readers often fail to notice contradictions in expository text (c.f. Otero & Kintsch, 1992), in our study we found significant slowdowns on inconsistent conclusions when students had all the required information to make the inference.

Reading Times (ms)	premise	neutral
Consistent Conclusion	1967	2025

Inconsistent Conclusion	2405*	2064

This indicates that when readers are given all necessary information in scientific passages, they are able to compute inferences. Singer has found similar results with contradiction experiments as well.

Experiments 3 & 4

Why were readers in this experiment able to recognize contradictory statements so well, while other studies have failed to find contradiction effects? Maybe it was the presentation of all of the relevant information in close proximity that promoted the generation of the inferences. In Experiment 3, filler material was inserted between the critical pieces of information in the passages, and the slowdown was lessened somewhat but was still significant.

Reading Times (ms)	premise	neutral
Consistent	2104	2147
Inconsistent	2346*	2174

In Experiment 4, we investigated two ways of distancing the critical information. The results indicated that there was a significant slowdown on contradictory information in both filler type conditions, although there tended to be less of a slowdown when the filler was placed between the premises and the conclusion, than when the filler was placed between the two premises.

	PPC		PPC	
	Premise	neutral	premise	neutral
Consistent	2370	2480	2100	2350
Inconsistent	2610*	2450	2560*	2300

This suggests that readers may be less likely to compute an inference when a conclusion does not immediately follow its premises. Perhaps inferences are only drawn upon encountering a conclusion, and perhaps only when both premises are still active. The results do seem to be converging on the conclusion that inferences can be drawn from scientific text when all the necessary information is available and active. One problem with all of the studies that have been done, however, is that we cannot determine exactly when the inference is computed. That is, does the reader compute the inference spontaneously as soon as all necessary information is available? Or is the generation of the inference prompted by the target probe, conclusion or contradiction? More work needs to be done to definitively address this question, and we are currently conducting an eyetracking experiment with these materials to this end. The study of inferences made from expository text is important because in many ways inference-making is at the heart of building a situation model, and ultimately, learning from text. With so much instruction occurring through scientific expository text, it is important to understand the conditions that promote inferences from scientific text specifically, and how they may differ from inferences from narrative text.

Previous Knowledge as an Influential Factor in the Clinical Diagnosis Inferences Generation:

An Expert and Novice study

J. A. Leon & O. Perez

Universidad Autonoma de Madrid (Spain)

joseantonio.leon@uam.es

Inference-making is a complex and necessary activity in comprehension processes that we use in daily life. An important factor in this context is previous knowledge which stands out as a decisive variable in the production of inferences. This is true to such an extent that we can affirm that without previous knowledge, inferences cannot be made (i.e. Leon, 1996). According to this assumption, we can add that an inference is made if we activate specific previous knowledge. In general, we consider that previous knowledge and the degree of expertise in a specific domain are influential in accelerating elaborative inference-making.

During the 70´s many studies examined the link between text comprehension and prior knowledge. The importance of this a priori organization and the ability to organize thoughts as they are encountered has been well documented in text comprehension literature (e.g. Anderson & Pichert, 1978; Bransford & Johnson, 1972; Fincher-Kiefer, 1992, 1993). Other research provides clear examples of how domain-specific knowledge influences reading (e.g. Chiesi, Spilich & Voss, 1979; Leon & Carretero, 1992; Spilich, Vesonder et al., 1979). Studies have documented that the knowledge of experts is organized differently from the knowledge of non-experts (Adelson, 1984; Chi, Glaser & Rees, 1982; Vonk & Norman, 1992). The knowledge of experts is organized in higher order knowledge structures, at a more abstract level, and according to general categories, laws and principles. These features are expected to accelerate the inferences that are product of dependent-specific knowledge during reading. On the other hand, many studies have focused on the role of inferences in comprehension, but concentrating mainly on narrative discourse (e.g. Graesser & Bower, 1990; Magliano & Graesser, 1991; McKoon & Ratcliff, 1992; Murray et al., 1993; Singer, 1988; Trabasso & Suh, 1993; van den Broek, 1990; van den Broek & Lorch, 1993). We therefore have investigated texts in clinical psychology in order to broaden our perception and understanding of inferences beyond narrative about everyday experiences.

In three experiments, we tested how domain-related knowledge (in the clinical psychology domain) influences clinical diagnosis inference generation during text comprehension. A clinical diagnosis inference is a particular kind of inference that is produced in a specific knowledge domain (clinical psychology and psychiatry). The activation of this inference involves a process of categorizing a given presentation of signs, symptoms, and behaviors shown by a patient in a particular context (clinical traits), and in a particular genre of text (clinical descriptions). Clinical diagnosis inferences can be activated when we ask what is the matter with him/her? For example, reading the sentence " Helen regularly experienced the urge to walk into department stores " followed by a second " She felt an irrepressible impulse to steal an article of clothing ", could trigger inferences such as a diagnosis of kleptomania. Clinical diagnosis is considered here as a kind of elaborative inference that could be included as an explanatory trait inference.

In this study, we have followed the three-pronged method (Magliano et al., 1991, Fincher-Kiefer, 1993; Trabasso & Suh, 1993). Firstly, the probability of generating clinical diagnosis inferences was analyzed by asking the subjects to think aloud (Trabasso & Suh, 1993; Trabasso & Magliano,

1996; Leon, van den Broek & Escudero, 1998). Secondly, we considered two theories (constructionist and minimalist) and their predictions about the status of clinical diagnosis inferences. Thirdly, we used lexical decision tasks and reading times as experimental measures to test the predictions. In the first experiment, we carried out a lexical decision task with a 500-ms SOA interval between the end of a sentence and the target word. The results showed that the expert group was significantly faster than the group of novices regarding the RD target condition. The mean obtained by them was around 750 ms. In the second experiment, self-paced reading times were recorded for some sentences (neutral, critical, postcritical 1 and postcritical 2) from each text. The data obtained by this experiment suggested that experts and novices showed different behavior, generating clinical diagnosis inferences at different times. For the experts, the inference-making process seemed to take place in its entirety when the critical sentence was read, while the novices generated inferences in the next sentence, confirming the results of the first experiment. Finally, in the third experiment we compared reading times for clinical and narrative texts, finding out that the differences between groups did not appear regarding to narrative texts and predictive inferences, which did not require any domain-specific knowledge. These results showed that the significant differences between experts and novices for clinical texts are caused by previous knowledge rather than by reading skills or general intelligence.

An important conclusion of this work is that the experts´ previous knowledge could be shown to be a decisive factor not only in the encoding and generation of the inference itself but also in determining when the clinical diagnosis is generated. Previous knowledge seems to accelerate the activation of inferences to such an extent that it can transform an inference originally considered off-line into an on-line one. The results presented here can also be related to the minimalist hypothesis (McKoon & Ratcliff, 1992, 1995) and to the constructionist theory (Graesser et al., 1994). Both theories indicate that, in general, elaborative inferences are unlikely to be drawn on-line unless they are supported by easily available or well-known information, are highly constrained by the context, and have few or any alternative outcomes. We agree with the assumption that the focus on all-or-none is insufficient to capture the fluidity of the processing and representation of even minimally complex texts (Gerrig & McKoon, 1998, p. 68). This conclusion coincides with the idea that on-line versus off-line as well as automatic versus strategic taxonomies are located in a continuum in which the inference can be activated rather than in an all-or nothing position as different research has shown (e.g., Gernsbacher, 1990; Graesser, Singer & Trabasso, 1994; Kintsch, 1988; Sharkey & Sharkey, 1992; van den Broek et al., 1999). This study suggests that previous knowledge is an influential factor in order to determine the possible location of the inference in this continuum.

The differences between experts and novices can be explained in terms of different mental representations, and the causal field from the reader's perspective. The well-known information on the part of the experts could reflect more elaborate and more sophisticated mental models that permit them to reduce the number of possible alternative consequences (Graesser et al., 1994). It is likely that the experts, unlike the novices, know how to contextualize information more quickly within a clinical context and that they make faster associations between the symptoms and the diagnosis. Our conclusions match those of previous studies (i.e., Patel & Groen, 1986), in which experts in a certain domain managed to eliminate irrelevant information by creating well-made partial situation models with highly discriminative features. On the other hand, novices generate both kinds of information -relevant and irrelevant- with evidence of the distinctive novice and intermediate peak. Novices create hypotheses with little features overlap (Patel & Groen, 1991).

Integration of Domain Knowledge from an Outline and a Target Text: Effects of Expertise, and Semantic Information

I. Tapiero & G. Molinari

University of Lyon 2 (France)

Isabelle.Tapiero@univ-lyon2.fr

The goal of this study was to investigate with high and low knowledgeable subjects the effect of the nature (states or events) of prior relevant knowledge on the learning of a scientific domain (the neuron). Prior to the reading of a target text, subjects were provided with an outline that referred either to the same or to a different category of semantic information (states or events) stated subsequently in the target text. We tested how the consistency or inconsistency between the type of knowledge provided in the outline (states and events) and the target text affected subjects' performances on the different levels of text representation (surface, semantic, situational).

Method

Eighty subjects (University of Lyon 2) participated in this experiment. According to their correct responses to preliminary questions about 4 notions related to the neuron (membrane potential, action potential, repolarization and period of non-excitability), subjects were set into two groups of expertise (high or low knowledgeable). Then, each group was assigned to the reading of one of five outlines designed to provide subjects with superordinate knowledge about the 4 notions described later in the target text. Each outline contained 4 paragraphs of 4 sentences each, one paragraph per notion and each sentence referred either to a state or to an event. In the first outline, the domain was described using more states than events whereas in the second one, the reversed principle was used. In the third outline, only "event" sentences were provided whereas in the fourth one, only "state" sentences were provided. Thus, although these four outlines differed in the category of information provided, they were informationally equivalent and were thematically related to the content of the target texts. The fifth outline dealt with information non-thematically related to the target texts and serves as a control. After studying one of the five outlines, subjects had either to summarize what they just read (summary condition) or to illustrate it on a schema (schema condition). This task was used to induce subjects to reinforce their semantic (summary) or situational (schema) representation of the outlines. Then, all subjects had to read a target text designed to provide specific knowledge about the 4 notions related to the neuron. As for the outline, the target texts contained 4 paragraphs of 4 sentences each, one paragraph per notion. Two different target texts were constructed, depending on the category of information the texts foregrounded (states or events). Thus, within each outline group, half of the subjects had to study a text that contained more states than events (" states " text), while the other half had to study a text that contained more events than states (" events " text). Memory for text and comprehension were tested via a recognition task with verbatim, paraphrases, inferences and domain-related distractor sentences. For each level, state and event category of test items were provided.

Our main assumptions were the following: According to what we observed in a previous experiment (Tapiero & Molinari, 1998), learning from text should result in building a coherent network structure in which " state " concepts are superordinated nodes directly linked to knowledge from the others semantic categories (events). Thus, in the present study, we expected a main effect of Category (states and events) on the memory representation (H1). At the situational level, states should be better represented than events, whereas less difference between these two categories should be observed at the two other levels (surface, semantic). We also expected an Outline by Text interaction (H2). According to Mannes (1994), subjects assigned to an outline that use the same category to describe the domain as does the target text (the SAME outline) should construct a more appropriate semantic representation than subjects assigned to an outline that use a different category (the DIFFERENT outline). On the other hand, subjects assigned to the DIFFERENT outline should construct a more integrated situation model than subjects assigned to the SAME outline. Moreover, we expected an effect of the type of tasks subjects had to perform (summary or schema) on the Outline by Text interaction (H3). Finally, we predicted a three-way interaction between Outline, Text and Expertise (H4). According to Mannes, Kintsch, Songer, & Kintsch (1996), we predicted a greater learning for experts assigned to the

DIFFERENT outline than for those assigned to the SAME outline. We expected the reverse pattern for beginners.

Results

Reading Time. As expected (H1), the " states " text required longer processing than did the "events" text. This effect could be explained by the fact that subjects assigned to the reading of the "states" text equally processed the two categories of semantic information (states and events), whereas subjects assigned to the " events " text focused preferentially on events. Moreover, our findings indicated that experts judged the events as being more important than the states to construct their mental representation of the text. On the other hand, beginners showed no difference in reading times for the " state " and the " event " sentences.

Recognition task (correct responses). As expected (H1), we found a differential effect of the two semantic categories (states and events) on the memory representation at the three levels of representation. At the surface level, states were better represented than events, while the reverse pattern was observed at the semantic and the situational levels. In addition, there was a Category by Expertise interaction. At the situational level, states were better represented by experts than by beginners, while no difference between the two knowledgeable groups was observed for the memory representation of events. Moreover, our findings indicated that memory for events was strongest when they were described in a text with more " state " sentences than " event " sentences. We assumed that the integration of the " event " information in memory would be facilitated when readers have more opportunity to relate them to the " state " informations in the text. Finally, there was a Category by Task interaction. Subjects induced to reinforce their semantic representation of the outline (summary condition) constructed a more appropriate representation of events than those assigned to the schema condition. The reverse pattern was observed for states.

Recognition task (recognition times). According to our expectations (H2), we showed a difference between the two outline conditions (SAME versus DIFFERENT) concerning the recognition time of inferences. This result indicated that readers who studied the SAME outline developped a different situation model regarding the category of semantic information foregrounded (states or events) than those who read the DIFFERENT outline. The longer recognition time for inferences observed in the DIFFERENT outline condition can be explained by the fact that when the outline and the target text referred to a different semantic category, it should promote more associations between propositions from the two contexts. Finally, an interesting result showed that beginners induced to reinforce their semantic representation of the outline (summary condition) spent less time to verify inferences than those assigned to the schema condition. At the opposite, experts constructed a more appropriate representation of the target text when they had to reinforce their situational representation of the outline (schema condition).

Complex Goal Structures and Narrative Comprehension

M. Singer & E. Richards

University of Manitoba (Canada)

msinger@Ms.Umanitoba.CA

This study inspected people's sensitivity to complex goal structure in the comprehension of stories. In contrast with the simple goal structures of many previous investigations, the present stories described two characters' attempts to accomplish independent, conjunctive subgoals in order to achieve a joint main goal. For example, Pam and Greg might have the main goal of meeting for lunch. To achieve the goal, Greg might have to catch a particular bus and Pam might have to complete a report.

In Experiment 1, the narratives consisted of: Three sentences which identified the characters' joint main goal; four sentences that described one character's subgoal that either succeeded or failed; four sentences describing the other character's subgoal, which always succeeded; a two-sentence target region in which the second character attempted to execute the main goal; and a two-sentence conclusion. Sensitivity to the conjunctive goal structure would be revealed by reading time that varied with the success of subgoal 1. Furthermore, to distinguish between constructionist and memory-based text processing analyses, the target region bore no surface nor semantic overlap with the four-sentence subgoal 1 region. Accordingly, a constructionist "consolidation" hypothesis (e.g., Long, Seely, & Oppy, Discourse Processes, 1996) predicted that reading time would be greater in the target region when the first subgoal had succeeded, and so permitted the interacting goals and subgoals to be integrated; than when the first subgoal failed. The memory-based analysis, in contrast, predicted no difference in target reading time. The results of Experiment 1 revealed longer Succeed than Fail reading times at the second target sentence, supporting the consolidation hypothesis.

In Experiment 2, both the success and the position (first or second in the story) of one of the subgoals were varied. The consolidation hypothesis predicted that, when the varied subgoal immediately preceded the target region, the longer reading time of the Succeed condition would be abolished: With the varied subgoal in this position, it would reside in working memory at the target in both conditions, and be equally available to participate in processing. The reading times replicated Experiment 1 when the varied subgoal was first, and supported the consolidation hypothesis prediction when the varied subgoal was second.

Experiment 3 directly assessed the consolidation hypothesis prediction that, with the varied subgoal first in the passage, it would be reinstated to working memory at the target in the Succeed condition but not the Fail condition. Consistent with this prediction, recognition time for a word representing the varied subgoal (e.g., BUS) was shorter at the target region in the Succeed than the Fail condition. In contrast, the memory-based analysis predicted, in view of the construction of the passages, no difference in subgoal-word recognition time as a function of subgoal success. The main conclusions were (a) that readers can represent relatively complex goal structures that are communicated by narratives, and (b) that these findings favour a consolidation hypothesis that is consistent with the constructionist comprehension framework over memory-based text processing.

Information Structure in the Processing of Sentences in Text

W. Vonk*/**, W. M. Mak** & J. C. J. Hoeks***

*Max Planck Institute for Psycholinguistics (The Netherlands)

**University of Nijmegen (The Netherlands)

***University of Groningen (The Netherlands)

Wietske.Vonk@mpi.nl

Processing the sentences of a text is a complex process, that involves the integration of different types of information, such as the syntactic category and meaning of the words, the syntactic and thematic structure of the sentence, and the preceding discourse, including pragmatic context. According to one class of syntactic theories, readers will initially assign the simplest possible syntactic structure in building up the representation of the sentence. These theories state that the sentence is initially parsed with help of the information of the syntactic category of the words only. Only later in the process, semantic and pragmatic factors play a role that eventually force the reader to repair faulty assigned structures.

In this paper we will show for some classes of sentences that processing these sentences in isolation is in accordance with predictions of the syntactic preference (Exp. 1 and Exp. 3), which for these sentences initially leads to a wrong interpretation and, consequently, to a long reading time. However, if these sentences are embedded in a proper context that manipulates the information structure, no sign of difficulty is encountered (Exp. 2 & Exp. 4). One can argue that the information structure of the text fragment overrules the syntactic preference, but we propose that the syntactic preference itself can be explained in terms of information structure.

In Experiment 1, coordinated structures were studied that are temporarily ambiguous between coordinated noun phrases and coordinated sentences, such as the sentence in example (1).

(1) The presenter thanked the author and the critic made the public a bow

According to Frazier's (1987) theory, readers assign the simplest syntactic structure to the incoming words, that is after encountering 'the critic' they assign the NP-coordination 'the author and the critic'. When a reader encounters the verb 'made', they must revise this initial analysis into a coordination of sentences, at some processing costs. In an eye movement study, these sentences were compared to control sentences that disambiguated the coordination towards a sentence coordination reading by inserting a comma before the 'and' in the coordination:

(2) The presenter thanked the author, and the critic made the public a bow.

The reading times revealed clear evidence for processing difficulty associated with the verb 'made' in sentence (1) compared to sentence (2), indicating the preference for an NP-coordination.

In Experiment 2 these sentences were embedded in two types of context, a biasing context, such as (3) and a neutral context, such as (4).

(3) Despite the late hour, the TV program for a life audience on art and culture always attracted a large viewing public. Because they had subjected the author to a heavy interview, the presenter and the critic were very satisfied.

(4) Despite the late hour, the TV program for a life audience on art and culture always attracted a large viewing public. This time the program ended with a heavy discussion on the latest best-seller.

The biasing context introduced explicitly a discourse topic consisting of two protagonists, 'the presenter' and 'the critic'. Therefore, it is likely that in the text that follows, two events will be described, one involving 'the presenter', and the other 'the critic'. So, when the biasing context

precedes sentence (1) it may be clear to the reader that 'the critic' is not part of a conjoint NP, but instead the topic of a new conjoint sentence. Indeed, in the eye movement measures there was no sign of processing difficulty in the biasing context associated with the verb 'made', compared to the control sentence, whereas in the neutral contexts the same result was obtained as in the sentences in isolation.

Comparable results are observed in the reading of relative clauses. In Experiment 3 relative clauses with a subject reading (SR-clauses), such as the sentence in (5) (SR reading: 'the jogger greeted the walkers') and relative clauses with an object reading (OR-clauses), such as the sentence in example (6) (OR reading: 'the walkers greeted the jogger') were contrasted. Note that these sentences are not disambiguated by word order in Dutch; they were disambiguated by the number information of the finite verb.

(5) Ongerust kijkt de hardloper, die de wandelaars in het park gegroet heeft, naar de donkere lucht. - 'Worried looks the jogger, that the walkers in the park greeted has, to the dark sky' - The jogger, who greeted the walkers, looks worried to the dark sky.

(6) Ongerust kijkt de hardloper, die de wandelaars in het park gegroet hebben, naar de donkere lucht. - 'Worried looks the jogger, that the walkers in the park greeted have, to the dark sky' - The jogger, whom the walkers greeted, looks worried to the dark sky.

Many experiments have shown for English, German, Dutch, etc. that subject relative clauses are preferred over object relative clauses. Most theories attribute this preference to the difference in syntactic structure between the SR and OR clauses, e.g. Frazier's Active Filler Strategy. In Experiment 3 this preference was again confirmed with eye movement measurements.

In Experiment 4 the referential expression of the protagonists was varied between a full NP and the indexical pronoun jullie, 'you' (plural), as in (7) and (8). Note that, as in English, this pronoun has no case marking in Dutch. The sentences are again disambiguated by the number information of the finite verb.

(7) Ongerust kijkt de hardloper, die jullie in het park gegroet heeft, naar de donkere lucht. - Worried looks the jogger, that you(pl) in the park greeted has, to the dark sky' - The jogger, who greeted you(pl), looks worried to the dark sky.

(8) Ongerust kijkt de hardloper, die jullie in het park gegroet hebben, naar de donkere lucht. - 'Worried looks the jogger, that you(pl) in the park greeted have, to the dark sky' - The jogger, whom you(pl) greeted, looks worried to the dark sky.

The indexical pronoun 'jullie' provides a discourse link to a highly salient entity. Therefore, readers will interpret it as the discourse topic, and assume that the pronoun is the subject of the relative clause. Indeed, the eye movement data associated with the finite verb were significantly longer in the SR clause (7) than in the OR clause (8). Thus these data indicated that not only the SR preference ('the jogger greeted you') disappeared, but even that it turned into an OR preference ('you greeted the jogger').

These two series of experiments show that the parsing preferences are not driven by a purely syntactic strategy. One can argue that the information structure of the text fragment overrules the syntactic preference, but we propose that the syntactic preference itself can be explained in terms of information structure. In the biasing context, the NP coordination preference was overruled by

the introduction of two protagonists that evoked two events. Two topics were indicated explicitly; therefore, the S-coordination was unproblematic. In the sentences in isolation, the NP-coordination preference can be explained by pragmatic factors of information structure, assuming that the readers obey the rule: "Don't assume a new topic unless indicated explicitly". Then the subject of the (first) finite verb is preferred as the topic of the sentence as a whole, and the critical NP must be part of an NP-coordination. In a similar way we propose that the SR-preference of the relative clauses is not a syntactic preference, but depends on pragmatic factors that determine which discourse entity is the most likely agent.

Gender Processing in French

S. Monnery, A. Seigneuric & D. Zagar

University of Bourgogne (France)

Alix.Seigneuric@u-bourgogne.fr

To understand a text, the reader has to process numerous morphosyntactic cues with high efficiency in order to build the syntactic structure of each sentence. Some of those cues are related to grammatical gender. Grammatical gender can be defined as a system in which the category to which a noun is assigned is reflected in the forms that are taken by other elements syntactically related to it (Comrie, 1999). Grammatical gender is a feature in many languages. It plays an important role as a device to establish local and global coherence in sentences and in discourse. For instance, gender marking conveys important information which helps the reader to determine the boundaries of Noun Phrase through agreement mechanism. In French, there are two gender classes, masculine and feminine. Nouns are assigned to genders by two principles. (1) Nouns are assigned to a gender according to their meaning, following a semantic principle (e.g., un garçon - a boy, MASC; une fille -a girl, FEM). (2) Nouns are arbitrary assigned in the case of inanimate nouns (e.g., un bureau - a desk, MASC; une chaise – a chair, FEM). In this last case, gender can sometimes be guessed on the basis of formal indications like word endings. For instance, most nouns ending in " ette " are feminine (e.g., une raquette - a racket, FEM), while most nouns ending with " on " are masculine (e.g. un ballon - a ball, MASC), although there are exceptions (e.g. un squelette - a skeleton, MASC ; une prison - a prison, FEM). Within the Noun Phrase, gender indication is carried by articles and adjectives.

In recent years, several studies have investigated gender processing during language comprehension. The issues usually addressed were: 1- How gender information affects processing and 2- when gender information is taken into account during reading. Empirical evidence has been largely based on a gender priming situation to explore the effect of prior congruent or incongruent gender information (provided by an article or an adjective) on the processing of subsequent words. The behavioral studies have employed different paradigms: A lexical decision paradigm but also naming or gating tasks. Gender effects have been explored in different languages such as Serbo-Croatian (e.g. Gurjanov, Lukatela, Moskovljevic, Savic, & Turvey, 1985), French (e.g. Colé & Segui, 1994), German (Schmidt, 1986) or Italian (e.g. Bates, Devescovi, Hernandez & Pizzamiglio, 1996). Some results are now reported. It has been observed that a lexical decision response to a target word was faster when preceded by a congruent syntactic word context (e.g., whose planet) than when it was preceded by an incongruent syntactic word context (e.g., it planet) (Goodman, McClelland, & Gibbs, 1981; Seidenberg, Waters, Sanders, & Langer, 1984). The lexical decision task is assumed to reflect post lexical processes so that the grammatical context effect has generally been interpreted as reflecting an automatic post access checking of gender congruity between context and target words. In the specific case of morphosyntactic agreement (e.g.,

gender, number, and case), it has been shown an inhibitory effect of inflectional inconsistency using pseudowords as experimental context items (Gurjanov, et al., 1985 for Serbo-Croatian). In French, Grosjean, Dommergues, Cornu, Guillelmon, & Besson (1994) have investigated gender processing with auditory presentation of both primes and targets in a gating task and a lexical decision task. Facilitation effect was observed when a noun target was preceded by a correctly gender-marked prime (article plus adjective) compared to a gender neutral base line (homophonous adjective only). Employing visual presentation, Colé and Segui (1994) have demonstrated that a lexical decision response to a target word was faster when this word was preceded by a word whose gender and number were congruent than when it was preceded by a word that disagreed in gender and/or number. This effect was observed even if the task did not require such a grammatical processing, suggesting that this effect was an automatic one. The effects already observed are far from providing a clear picture: Inhibition effects of syntactic incongruency have often been reported while facilitation effects remain more variable. Faced with conflicting results, opposite theoretical views have been proposed as interpretative frameworks for gender priming. According to one of these views gender effects can be described as prelexical. This interactive view postulates an interaction between syntactic and lexical information allowing subjects to predict lexical candidates on the basis of gender information (e.g. Bates et al., 1996). An alternative view describes gender effects as postlexical. This modular theory holds that lexical access is independent of prior syntactic or semantic information (Tanenhaus & Lucas, 1987).

The aim of this research is to study the processing of gender during reading. All the experiments presented here rely on grammatical priming in a lexical decision task. We have conducted three sets of experiments in skilled adult reader in order 1/ to replicate the Colé and Segui's previous findings in French, 2/ to specify the linguistic factors involved in gender processing, and 3/ to study gender agreement in larger Noun Phrase (e.g., determiner + adjective + noun).

In the first experiment, we ran a lexical decision task with inanimate concrete words. Each target was preceded by a prime which disagreed or not in gender with the target. Three kinds of primes were used to manipulate the syntagmatic relationship between the prime and the target: 1/ determiner (e.g., le texte, which constitutes a complete NP); 2/ adjective (e.g., beau texte, which constitutes an incomplete NP); 3/ noun (e.g., soir texte, with no syntagmatic relationship). The results showed that lexical decision times were longer when targets were preceded by incongruent determiners compared to congruent ones. The grammatical incongruency effect was smaller when the prime was an adjective, and there was no effect when it was a noun. These data suggest that the effect of agreement congruency emerges only when there is a syntagmatic relationship between the words.

The second set of experiments was conducted with larger Noun Phrase to study gender agreement between the different elements within NP. In these experiments, targets were preceded by two primes, a determiner and an adjective (e.g., un petit bureau, a small desk). The primes could either agreed or disagreed in gender with the target. We have manipulated the number of gender cues available to establish gender agreement. In French, some determiners and adjectives carry no gender cues. They can be described as hermaphrodite words (e.g., the adjective triste - sad; the determiner votre - your). We had three priming conditions: 1/ both elements carried information of gender (e.g., un (MASC) petit (MASC) bureau (MASC)); 2/ only the first element (the determiner) carried an indication of gender, the adjective being an hermaphrodite word (e.g., le (MASC) triste (HERMA) bureau (MASC)); 3/ only the second element (the adjective) carried an indication of gender, the determiner being an hermaphrodite word (e.g., votre (HERMA) petit (MASC) bureau (MASC)). The results showed that the effect of incongruency is observed whatever the number of cues and the distance between gender cues and targets. This suggests that processing of gender is effective as soon as gender cues are available during reading. The results will be discussed in the light of the debate between modular versus probabilistic theories.

How to Succeed with Telephone Answering Machines: Leaving a Message from a Psycholinguistic Perspective

J. Grabowski

Heidelberg University of Education (Germany)

fpjg@rumms.uni-mannheim.de

Although telephone answering machines (TAMs) are in private and easily accessible use since quite a few years, there are still many people reporting to feel uncertain or unpleasant when connected with a TAM, or they find themselves simply uncapable of leaving their message, or totally refuse to do so. Since language production, however, serves the regulation of the human cognitive system, they often skip a chance for successful achieving a communicative goal although it would, in many cases, be perfectly possible even when the partner is replaced by its TAM. What are the mechanisms that are responsible for this difficulty? We assume that one major obstacle is located in the early planning processes of language production (as opposed to the mostly automatic subprocesses of formulation and phonetic or grapho-motoric execution). As with all planning processes, language production draws, on this stage, on the restricted ressources of the cognitive system. For many cases, tasks or situations, we have schematized procedures about how to plan an utterance, e.g. when giving route directions or telling fairy-tales, which lighten the cognitive load during our language production. Apparently, many people did not acquire an appropriate pattern order schema for TAM talk. (Remember how, similarly, people living some decades ago when the telephone was a new and unfamiliar device would have raised their voices according to the presumed spatial distance of their partner, thus deriving their " ability " to make a phone call from their competence of how to behave during vocal interaction in natural situations.) What they do, instead, is either to treat the TAM situation like a normal phone call, or to try to use their letter writing abilities by formulating something like a spoken letter. (Or, third, they fail from the very beginning.) Formulating a spoken letter will, however, also fail due to cognitive overload: In writing, you can stop the formulation and execution processes and allocate your cognitive ressources solely to planning processes. With TAMs, however, you must continuously produce language, because the machine will break the contact in absence of incoming signals. The cognitive costs of the diverse subprocesses of spoken and written language production will be explained in detail. We assume that the use of a TAM will succeed best when the caller employs his or telephone conversation knowledge. The turn-taking in telephone conversation has been comprehensively described and analyzed since Schegloff's ethnomethodological approaches. Most important in this context is that turn-taking on the phone is described in terms of functional units (e.g., summons-answer, topic introduction, security turns for mutual understanding, good-bye turns etc.) as part of our shared social knowledge, particularly with respect to adjacency pairs. TAM conversation will succeed, if the caller can restructure the sequence of turns in phone conversation (which he or she perfectly masters) to two multi-turn units, one provided by the owner's message and the other replied by the caller's message. In order to underpin this assumption, we tried to show that (and in which way) the composition of the owner's message has systematic influences on the composition of the caller's message. To that end, we will report on two field experiments.

The first experiment was designed in order to show that the probability of callers leaving their message versus hanging up does depend on the owner's message's composition: If the way you communicatively deal with TAMs relates to your competence for spoken interaction on the phone, then your behavior should -- at least to some degree -- depend on the communicative offers of the owner's announcement message. We placed an advertisement for an apartment in a local newspaper. (For most people, the search for an apartment with a cheap rent is an urgent affair.) The advertisement appeared on a Friday (the local day of apartment advertising), when a TAM was connected with the given phone number. Between 7.30 a.m. and 1.30 p.m., three different messages were used, which were randomly changed by every hour. The three messages varied from a functional matter-of-fact like version to a detailed and more personal message, according to the functional units they were composed of. Our TAM recorded the spoken messages and counted

the total number of calls as well. During the observation period, 113 calls came in. The callers' readiness to leave their message varied significantly across the message versions they encountered: From 65 per cent in the neutral version condition (i.e., 35 per cent hung up immediately) to 30 per cent in the personal version condition (i.e. 70 per cent communicative fails). Thus, the basic readiness (or ability) to leave a message on the TAM depends on the communicative offer provided by the owner's message which gives first evidence for the assumption that the use of TAMs is to do with spoken interaction on the phone.

The second experiment was designed in order to show that the composition of the owner's message does not only influence the probability of the callers leaving their messages, but also the structure of these messages. We placed TAMs in two lawyer offices for a period of six weeks, operating outside of office hours. Again, there were three versions of the announcement message, which were systematically varied with respect to the functional turn units they were composed of and which were systematically changed according to a randomized plan. A total of 325 calls came in; 58 callers (18 per cent) used the possibility of pursueing their communicative goal and leaving a message on the TAM. These 58 messages are further analyzed in terms of turns they contain. It turned out that the composition of the callers' messages varied across the three versions such that they specifically reacted to the structure of turns in the owner's message. Several turn categories were significantly more or less frequently produced, according to the presence or absence of particular turns in the owner's message.

We conclude that if somebody succeeds in pursueing his or her communicative goal even when connected to a TAM, he or she appears to be able to employ the phone conversation schema by restructuring it into two multi-unit turns. Then, he or she is also able to specifically react to the communicative offer of the TAM message, completing the offered turns with the composition of the own message. Although these are preliminary results of a first investigation into our approach, there appear to be consequences (a) for instruction manuals on how to operate your TAM successfully (in terms of the true facilitation of social interaction), (b) for possible methods of training people whose work is closely connected to the distribution of information through the phone, and (c) for the didactics of teaching students how to become acquainted with the new media of individual communication.

Genre Change in the Historical Development of Sales Invitations

Y. Zhu

Institute of Technology, Auckland (U.S.A.)

Zyunxia@unitec.ac.nz

This paper sets out to examine genre evolution exhibited in the historical development of sales invitations used in mainland China in the Pre-Reform (1949-78) and Reform (1978 to the present) Periods. Here sales invitations refer to letters to invite the reader to sales exhibitions. The study of professional genre has been an area of research interest and relevant research can be found in Swales (1990) and Bhatia (1993). However genre evolution is still a relative new area and systematic research needs to be conducted. This paper will promote the study of genre evolution by examining genre development of Chinese sales invitations, in particular since after the economic opening-up in 1978. A corpus of twenty authentic letters collected from mainland China is examined as a specific example to illustrate the genre change.

The theoretical framework utilised in examining genre change in sales invitations is based on genre analysis (Swales, 1990) and schema theory (Bartlett, 1932; Rumelhart, 1980). As a particular advantage, this approach makes it possible to examine genre evolution in relation to the changing social context. Out of a changing context different communicative needs may arise. These needs are important factors and conditions for bringing about a change in communicative purposes. Consequently, a genre may evolve in various levels of the text. In addition, this approach incorporates culture specific knowledge into world and formal schemata to examine genre change. A specific feature may be related to the three kinds of reader-writer relationship which include " xiaxing " (the superior writing to the subordinate), " pingxing " (equals writing to each other) and " shangxing " (the subordinate writing to the superior). The writer has to choose appropriate forms depending on different reader-writer relationships.

Sales invitations are often described in textbooks as a type of " liyi xin " (letters of etiquette), and are treated as a form of etiquette for building relationships among people. This genre is characterized by a polite register, and linguistic continuums of (1) Polite salutations, (2) Inviting behavior, and (3) Polite closing can be found in textbook examples. Based on these characteristics, a detailed discussion about the writing of this genre is given to both periods. The Pre- Reform Period is characterized by an emphasis on a planned economy. In the hierarchical organizational structure, sales invitations played an important role to invite the subordinate organisations to the exhibition for product distribution. First of all, invitations tend to be " xiaxing " as the reader is in a subordinate position. Secondly, being informative and using the appropriate forms to invite the subordinate reader are the most important tasks of the letter. In the continuums of the polite forms, it has been found that the forms tend to be the least polite and the level of respect is rather low. Sales elements are completely missing. The Reform Period represents a link to the market economy. Sales invitations are experiencing dramatic changes in purposes, content and language forms since 1978. As a consequence, sales promotion becomes a new and strong component in sales invitations. It has been found, in the detailed analysis of the twenty-letter corpus, that besides inviting the reader, a series of new purposes are ascertained in this period such as attracting the reader's attention and interest, giving positive appraisals of the exhibition and persuading the reader to attend the exhibition. These new purposes are closely related to the marketing AIDA model. In turn, these purposes are realized in the lower levels of the text. The twenty letters, on the whole, share the new structural moves of (1) Establishing credentials (2) Describing the exhibition and (3) Offering incentives. All these moves can be seen as effective strategies to attracting the reader's interest and persuade the reader to buy or sell the product. Equal competition in business in the Reform Period may contribute to the change in reader-writer relationship from " shangxing " to " pingxing ". This change leads to a preference of more polite and respectful forms in salutations, inviting and closings in the continuums as compared to the Pre-Reform Period.

As further evidence to substantiate the genre change in the above respects, a detailed analysis is given to the most successful letter in the corpus rated by Chinese managers and their comments are also incorporated in the discussion. This letter is composed of ten moves: (1) heading, (2) salutation, (3) introducing the exhibition, (4) inviting the reader, (5) describing the exhibition, (6) offering incentives, (7) inviting the reader again, (8) providing registration details, (9) polite closing, and (10) signature and date. From these moves it can be seen that this letter is quite representative of the genre change in sales invitation. First of all, it has the polite invitation writing format as indicated in the formulaic moves such as heading, salutation and closing. The high level of politeness and respect is also shown in repeating the inviting moves. In Chinese culture, it is considered polite and respectful to invite the guest more than once. Secondly, the promotional strategies are well applied in relevant moves such as introducing and describing the exhibition, offering incentives and providing registration details. These tendencies of genre change are also found to be the major theme in the managers, comments as summarized below:

(1). This letter provides a good example for writing sales invitations. The reader is sincerely invited and the exhibition is well advertised.

(2). This letter represents a sincere and formal register to invite the reader which s exactly what a sales invitation should be like. The writer uses honorific forms when referring to the reader. The writer invites the reader twice, which indicates the writer,s hospitable attitude.

These comments, to a large degree, reflect the professionals, knowledge structures and general expectations about genre change in sales invitations. The first comment gives a general impression about this letter as both an invitation and sales letter. The second comment stresses the formal and respectful register. These two objectives are also seen being effectively realized step by step in the various lower levels of the text, such as steps and linguistic forms. For example, specific honorifics are employed; various kinds of adjectives and typical Chinese four-character expression are used as an important form of advertising language.

To conclude, this paper highlighted the importance of identifying the change in communicative purposes as a starting point to examine genre change. It was found that the purposes of sales invitation developed tremendously from distributing wholesales to inviting the reader to buy or sell the products. As a result, new advertising elements were introduced into sales invitations. On the other hand, there is a strong tendency for sales invitations to evolve from a " xiaxing " to a " pingxing " genre. This change can be shown in the preference of a more polite and respectful register. The above findings may have implications for the study of genre change in general; and changes in social context can be examined as an external factor affecting genre change. They may also have implications for doing business with Chinese enterprises.

Semantic Network(s) for Words and Pictures : Toward an Organization in Terms of Situation

M. Cornuejols & J-P. Rossi

University of Paris XI (France)

mc@limsi.fr

This research analyses the nature and the organization of the semantic associations in memory for verbal and pictorial entities. Through six experiments of priming, we demonstrate that there is a specificity of the associative network for pictures compared with the one for words that explain the difference of priming effects in automatic conditions. The composition and the organization of pictorial and verbal associative networks are examined in terms of situational and categorical links. The present research is focused on the nature and the organization of semantic representations in memory, underlying the comprehension process, and their relationships with the modalities of access. The aim of the research is to clarify if the semantic network which constitutes the semantic memory is a common structure for words and pictures or if there is a specific semantic network for each category of input, specially words and pictures ? Many types of problems can be identified in this framework:

- The way of access to the semantic memory: What is the semantic representation activated by words and by pictures ? Is it the same semantic content activated by the presentation of a picture of an entitie and by the name of this entitie ?

- The nature of the representation: What are the organization principles of the semantic networks ? Is there any difference in the organization of semantic the verbal and pictorial associated networks ?

1/ How do words and pictures activate the semantic nets in automatic condition ? The six experiments related here have been conducted to explore these questions.

Experiment 1: Automatic priming with written-words, heard words and pictures

The aim of the first experiment is to clarify if pictures, written-words and heard-words could activate the same associates in automatic condition. The paradigm of semantic priming is used with the lexical decision task. The pairs of items (prime-target) were established by reference to the French verbal association table (Lieury et al, 1976). Three types of primes are used: Pictures (coloured drawings), written-words, heard-words. The words correspond to the designation of object represented by pictures. The targets are always written-words. The classical effect of facilitation is found for verbal entities: the lexical decision is shorter when the prime and the target are semantically linked. This effect of facilitation is equivalent for written-words and for heard-words. However, the effect of facilitation is not found when the prime is a picture.

2/ It could be explained by the short time of SOA ?

Experiment 2: Pictures priming with long SOA

The lack of semantic priming effect for pictures could be explained by the short time of processing in automatic conditions. To test this hypothesis the second experiment consists into the replay of the experiment 1 with different SOA (stimulus onset asynchrony: Time between the beginning of the two stimuli). The paradigm used is the semantic priming with lexical decision task. The results show that semantic priming effect resulting of the activation of word by picture becomes apparent only if the SOA is superior or equivalent to 350 ms. This SOA is generally considered in the litterature as allowing a controlled (versus automatic) process of the prime. The hypothesis is that this specific time is sufficient for picture naming and so the priming effect is due to word (name)-word (target) priming and not results from a direct picture priming.

Experiment 3: Implicit naming of pictures in object decision task

To test the hypothesis of an implicit denomination of picture that occure in long SOA that allows priming effect, the experiment 3 was conducted in using an equivalent of the lexical decision task. This task is an object decision task that consists into asking participants to judge the reality of an object represented by a picture. The variable was the number of syllables that composed the names of pictures used as primes. This experiment shows that the response time is dependant on the length of these names that is to say that pictures induce implicite naming. It is important to notice that implicit naming doesn't mean automatic naming as Fraisse (1964) has demonstated: Words can be read faster than objects can be named, indicating the higher availability of the verbal code in the former case. So, the implicit naming, in the framework of dual coding theory (Paivio, 91), could explain that the names of the objects represented by pictures, equivalent to the verbal entities used before, activate the associates of the verbal norm.

3/ Is there any possibility of semantic activation by words and pictures in short time conditions ? The question remains: What happens in automatic semantic priming condition ? Is there no possibility to activate verbal network by pictures ? In other words, is there a specific associative network for pictures ?

Experiment 4: Elaboration and comparison of pictures and words association norms

Another explanation of the results of Experiment 1 is to consider that the semantic network for pictures is different from the one for words. To test this hypothesis of a dual semantic network, we have constructed an association norm for pictures (Cornuéjols, Rossi, & Brunschweiler, 1999) as one already exists one for words (exp. 4). This association norm concerns 284 pictures of artificial and natural entities of the everyday life presented as inductive entities. For each picture, 137 participants have to generate a mental image of the first associative entitie that came into mind and to write down the name of this associative entitie. To compare the pictorial norm with the verbal norm, we have also constructed a verbal norm with the names of pictures. The main result is that the two associations tables differed in 64 % of responses. A typology is proposed in terms of situation and categorical links providing many elements of discussion concerning the organization of associative networks. The proportions of associations that contextualized the entity (the situation, the localisation, the temporal links) and associations that specified (the characteristics of entities, for example) or categorized (semantic category, instancies, ?) the entity is different for pictures and words norms (Cornuéjols, 1999). Those results clearly show that reading a word or seeing a picture doesn't evoke the same associated item.

Experiment 5: Priming effects with new norms (verbal and pictorial)

The aim of the experiment 5 is to verify that these associations allowed automatic semantic priming and so to validate the norms by a semantic priming experiment in which the pairs of "picture prime and word target" arise from the pictorial norm, as the "word prime - word target" pairs arise from the verbal norm. These data show that a semantic priming effect occurs with picture prime in automatic condition. That validates the two associations norms constituted in the next experiment.

Conclusion for these 5 experiments: Thus all of those experiments suggest that associative networks for pictures and for words are specific. Then the question concerns the relationships between them. Are there two different networks, one for words and one for pictorial entities, or is there a unique network with differential regions, one activated preferentially by pictures and one activated preferentially by words ? The experiment 6 is conducted to anwer this question.

Experiment 6: Semantic priming with cross modal conditions

This experiment is conducted with experimental conditions in which prime and target could be a picture or a written-word in combination with the use of pictorial or verbal norms. The results show that in the conditions of semantic associations determined in reference of the pictorial norm, the effect of facilitation is obtained when the prime is a picture or a word and the target is a word. The semantic priming effect occurres only if the target is a word. The fact that a word could be activated by word and by picture under certains conditions of associations is in favour of a unique network with differential regions of activations. This conclusion is supported by the fact that time of semantic priming effects are similar between picture prime and word prime condition. In contrast, no activation of a picture occurres when the prime is a word or a picture.

This study demonstrates that: (1) the word and the picture don't activate the same associated item in automatic conditions time. The constitution of associations norms and the validation by semantic priming experiments have confirmed this conclusion. (2) The cross-modalities semantic priming effects could occurre only if the target was a word and not when the target is a picture

whatever the associative norm used. These conclusions raise the problem of the relationships between the two types of associative networks. The analysis of the semantic networks of pictures and words tries to clarify the differential organization of pictorial and verbal semantic systems that could explain differential of activations. A typology of associative links is proposed. This analysis points out that associations are mainly organized in term of situation and topics, in reference to questions of situational model, rather than categorical or taxinomic links.

Spatial Arrangement in Tables: The Influence on Reader's Mental Representation

A. Pellegrin & M. Bétrancourt

I.N.R.I.A. (Rhône-Alpes, France)

Anne.Pellegrin@inrialpes.fr

Most technical and educational documents contain not only text information but also different types of graphics. A large part of these representations consists of tables. Yet, among the large body of studies involving graphics, very few were interested in tables. Tables lie at the center of a continuum that extends from pictures to words. From words, they inherit the attribute of abstraction; but like pictures they exploit spatial layout in a meaningful way (Winn, 1987; Winn & Holliday, 1982). Thus, tables have the characteristics of both texts and images, although in contrast to text, the logical and syntactic relations which exist between the concepts are represented spatially on the page rather than in the sentence form. Research has attested that graphics are processed following the direction of written language (Tversky et al.,1991; Winn, 1982; Krohn, 1983). We carried out two experiments in order to investigate how the spatial arrangement of information in tables affects readers' cognitive processes. The first experiment (Pellegrin & Bétrancourt, 2000) investigated the effect of arrangement (horizontal *vs.* vertical) on the strategies used to describe the table, either to have it reproduced, or to make the addressee understand. In the description for reproduction task, the proportion of participants who used a hierarchical strategy over a linear one was higher when the arrangement was expected (i.e. horizontal for implication and vertical for genealogy). In the description for understanding task, we found a higher level of description complexity with an expected arrangement. A second experiment involving the same relations showed that in a verification task, performance was higher using an expected arrangement over a non-expected arrangement. These findings showed that the processing of information in tables is affected by the spatial arrangement of elements, which depends on the type of semantic relation expressed.

With regard to text, it is generally assumed that text comprehension involves the construction of multiple mental representations. In this framework, a mental representation of the linguistic surface is first constructed, from which the reader extracts a propositional representation of the text semantic content, which is then used to elaborate a mental model of the described subjects matter (van Dijk & Kintsch, 1983). Schnotz (1999) assumed that similarly, graphics involve the construction of at least two levels of mental representations: A visual representation (surface level) and a mental model (abstract level). Pinker (1990) has claimed that processing of graphics involves cognitive schemata, so called graph schema, elaborated from everyday perception and general world knowledge. Graph schema help elaborating a mental model from graphical material. On the basis of our previous findings, we assume that graph schema drive also the processing of tables. We assume that graph schema for tables correspond to the expected arrangement evidenced in our previous experiments. If information is plotted in an expected way, the mapping between the tabulated information and the graph schema is straightforward. Otherwise, processing of the table is more costful since the reader has to reorganize the information in order to map with the schema. We carried out an experiment to assess the hypothesis that information in table is coded following the graph schema, regardless of the actual arrangement of information.

Method

Participants. Twenty participants were individually tested. All participants were native French-speakers.

Material. The material consisted of 12 tables displaying an implication relation. 6 tables used an Expected arrangement (cause plotted to the left of effect) and 6 tables used an Unexpected arrangement (cause plotted to the top of effect).

Procedure. Participants were tested individually and received the same explanations. The table were presented on the screen using Psyscope software. The procedure involved two phases. In the first phase, the participants were presented a series of 12 tables. They were asked to identify the relation displayed in the table and to describe it orally. The table remained on the screen, which does not require the subject to memorize the table. This first phase is thus directed towards comprehension and should lead the participants to elaborate a mental model of the table. In the second phase, the participants were presented 18 tables: The 12 tables seen in the previous phase and 6 new tables. They had to determine for each table if they had seen it previously or not.

Design. The design was a repeated factorial design using the arrangement of elements in the table in the study phase and in the test phase, as within-subjects factors.

Table 1: Expected and Unexpected Arrangement in the two phases

Study phase	6 tables in Expected Arrangement		6 tables in Unexpected Arrangement	
Test Phase	3 tables Expected A.	3 tables Unexpected A.	3 tables Unexpected A.	3 tables Expected A.
Conditions	EE	EU	UU	UE

Results

For both (EE and UU) conditions, there was no effect of the spatial arrangement of the elements in the table on the test phase. There was a difference between EU and UE. For the EU condition, most of the subjects rejected correctly the table presented in the test phase (Unexpected arrangement). On the other hand, for the UE condition, most subjects did not reject the table presented in the test phase (Usual arrangement), as they should have done. The results are consistent with the hypothesis that when presented with a table in an Unexpected arrangement, subjects reorganize information in the mental model according to the graph schema, that is the arrangement we called Expected.

Discussion

The research we carried out on tables (Pellegrin, 2000) shows that, similarly to text, tabulated information leads to the construction of an abstract mental representation, whose construction and reactivation is influenced by graph schema in long term memory. From a practical point of view, this research suggests helpful hints for the design of tables in educational or technical documents. The findings particularly apply to specify computerized systems in which tables can be automatically generated from plain text. Since it is well known that readers experience difficulties in understanding complex graphical representations (tables, flow charts, and so on), further research must be carried out to provide designers with guidelines to improve the cognitive relevance of graphics.

References

Pellegrin, A. (2000). Traitement cognitif des présentations graphiques synoptiques. Thèse de doctorat de l'Institut National Polytechnique de Grenoble, Mention Sciences Cognitives, mai 2000.

Pellegrin, A. & Bétrancourt, M. (2000). How can spatial arrangement in table improve readers cognitive processing ?. in Proceedings of Smart Graphics, AAAI 2000 Spring Symposium. (pp 156-161) March 20th-22nd 2000, Stanford, CA, USA..

Pinker, S. (1990) A theory of graph comprehension. In R. Freedle (Ed.) Artificial Intelligence and the future of testing (pp. 73 - 126). Hillsdale, NJ: Lawrence Erlbaum Associates

Schnotz, W. & Grzondziel, H. (1999). Individual and co-operative learning with interactive animated pictures. European Journal of Psychology of Education, 14 (2), 245 - 265.

Tversky, B., Kugelmass, S. & Winter, A. (1991). Cross-Cultural and Developmental Trends in Graphic Productions. Cognitive Psychology 23:515-557.

van Dijk, T.A. & Kintsch, W. (1983). Strategies of Discourse Comprehension. New York: Academic Press.

Winn, W. & Holliday, W. (1982). Design Priciples for Diagrams and Charts. In Jonassen, D. (Ed.), The technology of Text. (pp. 277-299) Vol. 1, Educational Technology Publications, Englewood Cliffs, New Jersey.

What Determines Deep Comprehension for Illustrated Text?

S. N. Whitten, S. Lu & A. C. Graesser

The University of Memphis (U.S.A.)

SWhitten2@compuserve.com

It is often necessary to understand how a mechanical device works by reading and comprehending an illustrated text, such as the devices in David MacCauly's (1988) The Way Things Work. This task

requires a variety of skills such as verbal ability, mechanical ability, working memory, and spatial ability. The purpose of this study was to determine what factors predict a deep level of understanding for illustrated texts about how everyday devices work, such as those in MacCauly's book. We selected 6 devices from David MacCauly's (1988) book: A dishwasher, car temperature gage, clutch, cylinder lock, toaster, and electric bell. Each of these devices included a graphic representation and a brief description of how the device works.

The goal of this study was to investigate what cognitive abilities predict a deep level of understanding of the device, not a superficial representation - such a list of parts associated with the device. By deep comprehension, we mean a representation of the device that includes the components, spatial arrangement of the components, the causal relationships among the components as the device works, and the mechanisms that explain each step in the causal chain of events (Graesser, Pomeroy, Craig, & Olde, 1999). In order to test the reader's deep comprehension of the device, two assumptions were made about deep comprehension. First, because a deep level of understanding requires an explanatory model of how the device works, understanding is best demonstrated when the device breaks down, not when it is operating smoothly. Second, it has been suggested that the quality of the questions generated by the reader indicates a deep level of understanding (Graesser et. al., 1999). In other words, understanding of the device is best demonstrated when the reader asks a good question about what may be wrong with the device. However, the Graesser et al., study never investigated the serial positions of the questions that were asked. This study will examine how cognitive abilities relate to the order of the questions generated by the comprehender. The goal of the current research was to identify the measures of individual differences that predict the quality of the first two questions asked as opposed to the quality of the remaining questions. The measures investigated were: gender, openness (or creativity), working memory, exposure to print, spatial ability, coding speed, technical ability (mechanical reasoning ability), verbal ability, and quantitative ability. Data were collected on all these variables after the experimental task of generating questions in the context of a breakdown scenario.

Method

Participants. For this study, 108 students at the University of Memphis participated for course credit.

Procedure. There were two phases of this experiment. In the first phase, participants were shown the devices one at a time. Then, they were given a situation in which the device broke down. The participants were asked to generate questions about what may have been wrong with the device. In the second phase, the tests of individual differences were administered. This included the ASVAB (Armed Services Vocational Aptitude Battery, Department of Defense, 1983), which measured verbal, technical quantitative, and coding speed; exposure to print, ART (Author Recognition Test, Stanovich & Cunningham, 1992); NEO personality inventory (Costa & McCrae, 1991), which included a measure of openness; Working memory span (LaPointe & Engle, 1990); and spatial reasoning (Bennet, Seashore, & Wesman, 1972).

Results

A separate regression analysis was performed on the 9 predictors for the total quality of question, the first two quality questions, and the remaining quality questions. For the total quality questions, the regression equation was significant, $F_{(3, 39)} = 7.0$, $p > .001$, $R^2 = .39$. It was found that technical ability predicted the overall quality of questions with a beta weight of .445. For the first two quality questions, the regression equation was also significant, $F_{(3,39)} = 3.4$, $p > .01$, $R^2 =$

.24. An analysis of the quality of the first two questions asked yielded the same result with a beta weight of .325.

For the remaining quality questions, the regression equation was significant, $F (3, 39) = 3.6$, $p > .01$, $R^2 = .25$. However, the quality of the remaining questions (omitting the first two) yielded a different result. For these remaining questions, working memory was a significant predictor with a beta weight of .241 and spatial ability was a significant predictor with a beta weight of .253.

Discussion

It seems that in understanding a description of a mechanical device, technical ability is initially the best predictor of deep comprehension. However, working memory and spatial ability predict deep comprehension after some time has elapsed.

References

Bennet, G. K Seashore, H. G.& Wesman, A. G. (1972). Differential aptitude test: Spatial relations, Form T. New York, Psychological Corporation.

Costa P.T., & McCrae, R.R. (1991). NEO: Five Factor Inventory. Odessa, FL: Psychological Assessment Resources.

Department of Defense, (1983). Armed Services Vocational Aptitude Battery, Form 12a. Washington, D.C., Department of Defense.

Graesser, A. C., Craig, S., Pomeroy, V., & Olde, B. (1999, April). Deep comprehension of illustrated texts in the context of a breakdown scenario. Invited symposium on discourse comprehension at the meetings of the American Educational Research Association, Montreal, Canada.

LaPointe, L. B. & Engle, R. W. (1990). Simple and complex word spans as measures of working memory capacity. Journal of Experimental Psychology: General, 64, 1118-1133.

Stanovich, K.E. & Cunningham, A.E. (1992). Studying the consequences of literacy within a literate society: The cognitive correlates of print exposure. Memory and Cognition, 20, 51-68.

Cognitive Maps Constructed from Text and Virtual Navigation: Effects of Modality and Prior Knowledge in the Representation of Spatial Environments

A. Tan & C. R. Fletcher

University of Minnesota (U.S.A.)

atan@cogsci.umn.edu

The results of previous studies show that readers can encode spatial information described in a text (e.g., Perrig & Kintsch, 1985). Attempts have been made to characterize the spatial situational models constructed from reading text. Investigations have focused on the spatial perspective maintained while reading a spatial description, the task demands that lead to the representation and inference of spatial relations, and whether situational models can represent metric, Euclidean properties.

In the area of environmental cognition, cognitive mapping theories postulate the spatial memories that people form of environments are highly distorted and are systematically influenced by Non-Euclidean factors (e.g., downtown vs. suburb). Many theories argue that people hierarchically organize spatial information in memory (Hirtle & Jonides, 1985; Kosslyn, Pick, & Fariello, 1974). These non-spatial components that lead to the chunking of the environment in memory exert an influence on the spatial judgments that people make. It biases their response about the locations and distances between objects, such that they overestimate distances between objects across regions and underestimate distances between objects within regions.

The current study assessed the organization of verbally presented spatial information in memory. To determine the nature of this organization, a direct comparison was made between the representations derived from text and those constructed from virtual navigation. We predicted, based on past research findings, the cognitive maps constructed from texts would show similar systematic errors to those acquired from virtual navigation of an environment. In addition, the impact of prior knowledge (i.e., spatial schemas) on the hierarchical organization of spatial memories was studied. Spatial schemas are expectations that people have for highly familiar, stereotypical environments (Brewer & Treyens, 1981). These expectations should further attenuate these distortions because of strong pre-existing associative connections in memory.

Materials and Procedure

Virtual Layouts. Two virtual environments were generated using Virtus Walkthrough Pro software. The two environments were identical to each other in every respect except in the objects placed inside them. The schema-consistent environments were filled with objects typically found in an apartment. In contrast, the schema-inconsistent environment contained objects that were not functionally related. A virtual tour was pre-recorded, such that the presentation order of the objects was the same across the two environments.

Text Layouts. Two texts, each 72 sentences long were constructed. The two texts described the same spatial information depicted in the virtual environments. The texts were exactly the same in every respect except in the sets of objects used (i.e., schema-consistent vs. schema-inconsistent). The texts each described an environment as if the reader was walking through it. The presentation order of objects was the same as the virtual tour. The descriptions used a mixture of route and survey perspective statements to describe the locations of objects. Precise distance information was not stated.

Spatial location judgments. A total of 13 spatial location judgments were constructed. The participants were asked to imagine themselves from a specific location and perspective. They were then probed for the relative location of an object from that imagined position. Participants were only required to provide a categorical spatial response (i.e., front, back, right, or left) to the question.

Distance judgments. A total of 84 triadic distance comparisons were constructed. Each triad consisted of three pairs of objects. Each pair in the triad represents the distance between two objects described or depicted in the environment. Participants were required to rank order within each triad the three pairs of objects from closest together to farthest apart.

The study was divided into two phases: A learning and testing phase. In the learning phase, participants were asked to watch a computer-generated virtual tour or read a text description of the same environment. Next, participants were asked the spatial location judgments. They were given feedback about their answer after each question. Participants, who incorrectly answered a question, were required to re-view the animation or re-read the texts. After reaching a criterion level of learning (13/13), participants moved on to the testing phase. Participants were given a list, containing the triadic distance judgments, and asked to rank order from closest to farthest the different pairs of objects.

Results and Discussion

This was a 2 x 2, between participants design with the modality of presentation (i.e., text vs. virtual navigation) and schema-consistency (i.e., schema-consistent vs. inconsistent) as the main variables. To assess the organization of the spatial representations formed, two derived estimates (i.e., rooms apart and actual distance dependent variables) based on the original distance judgments were analyzed. First, the rank order distances between pairs of objects were entered into a multi-dimensional scaling analysis to give a two-dimensional output. Absolute distances between object pairs were calculated on this output. These estimated distances were then entered into a regression analysis. Two variables, the actual distance between objects and the number of rooms separating the two objects, were entered into a regression analysis. The regression coefficients computed reflect the fidelity of the representations formed. Higher values on the "actual distance" coefficient indicate greater improvement in the accuracy of the distance estimations. In contrast, higher values on the "rooms apart" coefficient indicates greater degree of systematic error in the distance judgments, and thus more hierarchical clustering of the spatial information.

An analysis of the "actual distance" measure did not reveal any main effects of either modality of presentation or schema-consistency, or an interaction between the two variables. However, planned contrasts did identify significant differences between groups. Specifically, the analysis showed that participants who virtually navigated the schema-consistent environment were significantly more accurate in making distance judgments than either the participants who virtually navigated the schema-inconsistent environment or those who read a text about the schema-consistent environment. An analysis of the "rooms apart" measure also did not reveal any main effects of either modality of presentation or schema-consistency of the environment, or an interaction between the two variables. But planned contrasts did show significant differences between groups. Specifically participants who read about the schema-consistent environment showed greater evidence of hierarchical organization in their spatial memories than those who read about the schema-inconsistent environment. Moreover, this effect was significantly greater for participants who read rather than virtually navigated the schema-consistent environment.

The primary goal was to demonstrate the cognitive maps that people form from either reading about or navigating an environment are similar to one another in that they are hierarchically organized. If participants' cognitive maps had accurately represented Euclidean space then the values of the "rooms apart" regression coefficient should have been near zero. This was not the case. Moreover, the main effect of mode of presentation was not found suggesting a similar degree of hierarchical organization between readers and navigators. However, this effect was mitigated by the influence of schemas. Readers, who were given the "apartment-like" environment, showed significantly greater hierarchical organization in their cognitive maps than those who read about the "non-meaningful" environment. This finding is consistent with the idea that non-spatial factors such as function can influence the organization of spatial information in memory.

Interpersonal Expectations in Comprehension of Visual Narratives

E. Gámez, H. Marrero & J. M. Díaz

University of La Laguna (Spain)

egamez@ull.es

The so-called "constructionist viewpoint", (Graesser, Singer & Trabasso, 1994) proposes a model of reader who, to comprehend, has to construct a deep representation of the discourse, that is, a representation which includes the underlying causal structure of the story. This implies therefore an active reader who constructs a meaningful situation model that is compatible with the text, and for which it is necessary to make inferences explaining the reasons for the actions, events and states mentioned therein. Of the group of possible inferences produced in the course of reading, the above authors give central importance to inferences about the goals of the characters in the stories.(Huitema, Dopkins, Klin & Myers, 1993; Long & Golding, 1993; Long, Golding & Graesser, 1992; Lutz & Radvansky, 1997; Magliano, Bagget, Johnson & Graesser, 1993; Suh & Trabasso, 1993; Trabasso, van den Broek & Suh, 1989; Trabasso, 1991). However, research reviewed from the constructionist point of view (Graesser et al., 1994; Singer et al., 1994) usually studies the role of "goals" and "motives" as functional categories in the text. That is to say, it studies the links between goals, motives, actions and states without explicitly determining exactly the nature of these goals or motives. Since the pioneering research of Schank and Abelson (1977) and Schank (1982), few authors have examined either the connections between goals, motives and behaviours within specific areas of experience or the source of world knowledge or "themes" from which we can infer goals which are not made explicit in the narratives (Dyer, 1983; McKoon, Ratcliff & Seifert, 1989; Seifert, Abelson & McKoon, 1986; Seifert, Abelson, McKoon & Ratcliff, 1986; Seifert, Dyer & Black, 1986).

In this poster we investigate the role of a type of motivational knowledge in the process of comprehension of character's action in the interpersonal realm. Specifically we suggest that the affective characters' relations represent a type of motivational knowledge that readers identify in the course of comprehension in order to anticipate the goals and actions of a character in the context of a situation where one character explicitly or implicitly requests help from another. In previous research we used texts where an interpersonal episode was narrated and we found evidence for the activation of interpersonal expectations (goals, actions and character's relations), (Gámez, 1995; Marrero, Gámez, Espino & Castillo, 1995; Gámez & Marrero, 1999). In the experiments presented here, we wished to explore previous predictions in the context of visual narratives. These visual narratives included seven cartoons per story and each of them expressed a scene of the episode narrated. The first cartoon expressed the relationship between the characters which could be close or distant. When the cartoon expressed a close relationship we showed characters talking together, sharing some objects, and so on. When the cartoon expressed a distant relationship we showed the characters without interaction. In the second, third and fourth cartoons we showed the causes of the interaction between characters (i.e. a character had a problem, a third character attacked another, a character needed something that the other had, and so on). In the fifth cartoon the "helping" situation appeared, this was the moment when one character requested help from the other. The sixth cartoon showed a positive (one character helped the other) or a negative action (one character did not help the other). The last and seventh cartoon ended the sequence with a panorama of the scenario in which the story had occurred.

In Experiment 1 we recorded the time that participants spent viewing the sixth cartoon of each sequence. This target cartoon showed an action that could be congruent or incongruent with the first-seen cartoon, and expressed an affective characters' relation that could be close or distant. Our predictions in this experiment were that when participants viewed the first cartoon with a close relationship and the sixth cartoon with a positive action they would view the target cartoon in significantly less time than in the condition with a distant relationship. This would suggest that when participants viewed a close relationships they activated an interpersonal motive related with an action congruent with this motive. The results only showed a main effect of the Type of Action (positive or negative), $F_1(1, 47) = 17.43$, $p < 0.001$ and $F_2(1, 11) = 14.06$, $p < 0.003$. Participants viewed cartoons which expressed a positive action ($M = 2770$), in less time than cartoons with negative actions ($M = 3248$), independent of the previously perceived relationships. These results do not support the presence of viewers' expectations about specific action in the course of comprehension coherent with a motive of maintaining the relationship explicit in the story. A possible explanation could be that the cartoon that expressed the distant condition only reflected a more ambiguous cartoon than the close relationships. This means that participants understood that the distant relationship was only a less close but not a different valence relationship

In Experiment 2 we substituted the "viewing times", for latency in naming a word representing the interpersonal goal of each story. The interpersonal goal appeared between the fifth and sixth cartoon, because the fifth cartoon expressed the "helping situation" where one of the characters asked for help or support from the other character. Participants could view stories with an explicit relationship (close or distant) in the first cartoon and the interpersonal goal (positive or negative). Now, our predictions were that there would be shorter naming latencies of the positive goal when participants viewed a close relationship in the first-seen cartoon. The first analysis of data from experiment 2 showed a tendency in the interaction between the Type of Relationship x Type of goal $F_1(1, 39) = 1.23$, $p = 0.27$, $F_2(1, 11) = 3.50$, $p = 0.08$. The effect of Type of Relationships and Type of goal was also significant [$F_1(1, 39) = 4.46$, $p < 0.04$, $F_2(1, 11) = 2.25$, $p = 0.16$ and $F_1(1, 39) = 12.66$, $p < 0.001$, $F_2(1, 11) = 6.03$, $p = 0.03$. In a second analysis we eliminated 3 of the 12 stories used in the experiments because we later detected some inconsistencies in the presentation of the stories. In this new analysis there was a significant interaction of Type of Relationships and Type of goal $F_2(1, 8) = 5.58$, $p = 0.04$. We found shorter latencies of the positive goal in the close condition ($M = 816$) than in the distant condition ($M = 894$). These results support the presence of a certain expectation about a general goal such as: To help, to defend, etc in the course of comprehension of interpersonal episodes where the context is of helping situation.

The results of Experiments 1 and 2 do not support a strong version of the "constructionist view" where the principle of search-after-meaning requires the maintaining of an expectation in the course of comprehension. We found some evidences related to participant activation of an expectation of a goal in the locus of a "helping situation" when there was a close relationship. However, we need to explore what happen with the activation of a specific action. Perhaps the viewer only has an abstract expectation? We would like to stress that the nature of cartoons implies a more ambiguous interpretation from the observer/participant than when the episode is narrated using a written text.

Computerized Document Search by Young Learners: Effects of Conceptual Organization in Memory on Search Terms Production

J. Dinet*, J-F. Rouet* & J-M. Passerault**

University of Poitiers * (France)

C.R.D.P. of Poitou-Charentes (France)

jerome.dinet@mshs.univ-poitiers.fr

Students are often asked to find documents (e.g. in books, CD-ROM, Web pages, etc.) in order to acquire information on a topic of pedagogical interest. Finding documents requires the student to interact with "indirect databases", e.g. library catalogs and bibliographic databases. These indirect databases only give the user a pointer to a computerized document space which is a list of document references. In most cases, and especially in the case of computerized search tools, the student must produce keywords in order to receive a list of document references (Bares, 1984; Sutter, 1984). Thus the production of relevant search terms is essential in finding appropriate documents. In the elementary grades document search is generally performed only after reading general information about the topic, through e.g., a passage in a textbook. The preparatory reading helps the student activate relevant knowledge and thus increases the likelihood that he or she will generate appropriate keywords during the search. In this paper we examine the influence of various preparatory activities on the amount and relevance of keywords generated by elementary grade students as part of a document search task. From a cognitive point of view, the influence of prior reading on the efficiency of a search task may be interpreted in terms of patterns of concept activation. According to the Landscape Model (van den Broek, Risden, Fletcher & Thurlow, 1996; van den Broek, Young, Tzeng & Linderholm, 1998), as well as in other recent models, reading is conceptualized as a cyclical and dynamic process. The mental representation is updated in the course of processing of the text. The patterns of activation across reading cycles form the basis of the memory representation that the reader constructs from the text. Reading a passage of text prior to a search task may influence the "cohort" of concepts that will be readily available prior to the search time. Exactly which concepts will make up the cohort is a function of text and reader characteristics. There are two major techniques for modifying the conceptual organization in a mental representation of a reader: The text structure and the reader involvement in the task. Several studies have demonstrated that the surface structure of the reading material affects the mental representation of the reader. Structure may be provided through typographical, linguistic and structural cues like paragraphing, which guides the allocation of reader's attention during reading (Lorch & Lorch, 1985, 1996; Lorch, Lorch & Inman, 1993). In a recent study, Gaddy, van den Broek and Sung (in press) have demonstrated that the form of the reading material affect activation patterns and the construction of the final representation of an expository text. Other studies have demonstrated that promoting reader involvement affects the construction of the final representation of a text. Mayer (1984) identified three kinds of aids to text comprehension for young readers: 1) aids for organizing or for building coherent connections between main information, for example by structured notetaking, 2) aids for integrating or for building external connections, for example by helping to activate background knowledge, 3) aids for selecting principal ideas, for example by underlining main words or sentences in a text. Underlining was also used by Vezin (1975, 1978) to help young readers in text comprehension. Be as it may, structuring the materials and promoting active involvement may help the reader to conceptually organize his or her mental representation during reading. In accordance to the Landscape Model, we can suppose that structuring the materials and promoting active involvement influence the patterns of activation across reading cycles which form the basis of the memory representation that the reader constructs from the text (i.e., influence the "cohort activation"). What are the effects on the search terms production ?

Method

Participants. The participants were 87 pupils recruited from primary grades (mean = 9;8 years-old) in the educational district of Poitou-Charentes (France). All participants were French speakers and they had practiced computerized document search every school week for two years.

Procedure. First, the students were asked to read a 110-words text about medieval history. The text dealt with three main topics: Gothic art, knights' crusades, and knights' armors. Each student was randomly assigned to one of three study conditions. In the Control condition the topics were presented in a same paragraph, without structural cues. In the Paragraphing condition the topics were presented in three paragraphs (one topic per paragraph). In the Underlining condition the topics were presented in a same paragraph (like in Control condition) but participants were asked to underline sentences which were semantically related, using different color pencils. Immediately after reading the text, the pupils were asked to perform a computerized search task on the same topic. The task required the pupils to produce search terms about gothic art, crusades and armors.

Medieval history knowledge was assessed by a twelve-item multiple-choice questions. There were four questions for each of the three topics of interest. For each question there were five possible answers including 'I don't know'. The participants were asked to check the right answer for each question.

Scoring and Results

Quantitative aspects. We computed the raw number of search term produced, as well as the number of search term different from the assigned topics (i.e., gothic art, crusades, armors) that we called inferential search terms. Pupils in the Control condition wrote significantly less search terms than pupils in the Paragraphing or Underlining conditions for each topic. In addition, a smaller number of pupils did produce inferential search terms in the Control condition. No other significant result was observed.

Qualitative aspects. For each topic, three computerized document search experts were asked to evaluate the relevance of each inferential search term. We noted the four inferential search terms which were the most relevant for each topic. We computed the number of these relevant inferential search terms for each condition. We found that the number increase significantly with conceptual organization for each topic. In other words quantitative and qualitative aspects of the search terms were affected by conceptual organization whatever level of topic knowledge.

Discussion

The differences across conditions may be interpreted in the terms of the "mental model strategy" proposed by Mayer, Dyck & Cook (1984): A reader uses a "mental model strategy" when he or she attempts to build a coherent mental model of the text, emphasizing the key nodes and links between them. In contrast a reader uses a "default strategy" when he or she can't emphasize the links between the key nodes. Each paragraph of the text used in our study consisted of a set of nodes, with each node corresponding to a concept related to the paragraph's main topic (for example, "shield" is a specific concept of the "knight's armors"). In addition the three main concepts are directly related to the general concept of medieval history. It is likely that the pupils in the Paragraphing and Underlining conditions were induced into a "mental model strategy" because these preparatory conditions helped them to conceptually organize their mental representation, i.e., their cohort activation (Gaddy, van den Broek & Sung, in press). In contrast pupils used a "default strategy" in the Control condition because they didn't have any help for structuring their mental representation. These results suggest that computerized document search by children may be successfully guided through appropriate preparatory conditions.

Referential Information and Composition Strategy: Differing Influences on Temporal Variables

C. Dansac

University of Poitiers (France)

Christophe.Dansac@mshs.univ-poitiers.fr

The research reported here explores how the writing process of organizing content (Hayes & Flower, 1980) deals with domain knowledge and more specifically with knowledge structure. Referential information to be communicated within a composition is assumed to have three essential characteristics resulting from the storage of knowledge within long-term memory: Accessibility, hierarchical integration and intrinsic linearity. We assume that these characteristics influence the process of content organization. Moreover, it is assumed that this influence differs according to the composition strategy writers engage in. Bereiter and Scardamalia (1987) propose that text production can proceed either by a knowledge telling strategy or by a knowledge transforming strategy. In the first strategy, "tellers" translate content as and when it is retrieved from long-term memory. In the second strategy, "transformers" enrich their text content problem space by a dialectic process that takes into account discourse constraints and rhetorical concerns. In our view, a writer using a knowledge transforming strategy should be more influenced by knowledge structure than a writer applying a knowledge telling strategy should. And the temporal parameters of written production should be affected by the properties of the to-be-communicated-content in differing ways according to the strategy the writer engages in.

Two experiments tested these hypotheses using an innovative paradigm. Having subjects writing a text about an imaged referent allowed us to reveal their composition strategy. Indeed, confronting the structure of information within the imaged referent with the structure of information within the text enables one to identify what particular strategy has been adopted. Also, in both experiments, a particular recording device was used to collect the productions. A digitizing pad recorded the handwritten production. The G-Studio software (Chesnet, Guillabert, & Espéret, 1994) operated the pad, controlling for the moments when the image was showed on the screen. The writer had to touch with the pen a particular zone of the tablet in order to have the image displayed. So the writer was unable to write and look at the image at the same time. This device permitted to exhibit the moments when referential information was accessed and processed. Those moments simulate accesses to LTM, and thus content retrieval, and their duration reveals the cost of the processes applying on the information retrieved.

The first experiment was conducted with 135 undergraduate students. We manipulated the hierarchical integration and the intrinsic linearity of the content to be communicated using fictitious geographic maps varying as regards the disposition of economic resources. The duration of the accesses to the image and their location within the text structure were analyzed. The results showed that manipulating the characteristics of the referent influences the temporal course of composition differently in tellers and in transformers. The way referential accesses are dispatched among the course of the activity is only related to strategy; how long referential accesses were was dependant both on their location within the text and on the characteristics of referential information.

Fifty adult subjects participated in the second experiment. A particular writing device was used in order to control for the moments when the writer reread their text. Subjects could not reread their text on-line, but could however consult their product in the same manner they consulted the map. Thus, this experiment also allowed to gain information about how rereading dispatched amongst the production. The frequencies of rereading phases paralleled the frequencies of image

consultation. The analysis of how accesses to referential information dispatched and how long they were confirmed some of the observations in the first experiment. However, the two composition strategies did not differ as clearly as they did in experiment one. The difference in the results between the two experiments are interpreted with regard to 1) the difference in the written products collected, 2) the writing environment.

The cyclic nature of accesses to referential information and its correspondence with the text structure observed in those two experiments parallels what has been observed in comprehension studies. This leads us to propose that the organizing process be fundamentally the same as the ones involved during text comprehension particularly when the reader has to reduce the information (see van Dijk & Kintsch, 1983). Organizing is thus considered as a process of deriving macropropositions from a set of propositions held together within working memory. The strategic aspect of organizing is accounted for by an active inhibition of a particular set of macropropositions in order to reach the elaboration of an alternative set of macropropositions: then organizing may be considered as a controlled process, relying on working memory. This view is compatible with other data obtained in production studies using secondary tasks interfering with the central executive component of working memory (Dansac & Passerault, 1998). The theoretical proposition that assimilate organizing to a comprehension process introduces some coherence across the various writing situations: the same processes are involved when the writer plans from head, and when the pieces of text or notes already produced are reread in order to find a structure. Extending the comprehension perspective to the whole production process allows envisaging how the early phase of content generation can be related with assignment reading. It should also allow exploring the initial phase of text planning within the framework of the construction-integration model (Kintsch, 1988, 1998).

References

Bereiter, C., & Scardamalia, M. (1987). The psychology of written composition. Hillsdale, New Jersey: Erlbaum.

Chesnet, D., Guillabert, F., & Espéret, E. (1994). G-Studio: un logiciel pour l'étude en temps réel des paramètres temporels de la production écrite. L'Année Psychologique, 94, 115-125.

Dansac, C., & Passerault, J.-M. (1998). Disrupting central executive or phonological loop: Effects on the writing process and on the writing product. In M.F. Crete & E. Espéret (Eds.) Proceedings of the SIG Writing conference 'Writing and Learning to write at the dawn of the 21st century'. Poitiers : MSHS.

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), Cognitive processes in writing (pp. 3-30). Hillsdale, New Jersey: Erlbaum.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: a Construction-Integration model. Psychological Review, 95(2), 163-182.

Kintsch, W. (1998). Comprehension: a paradigm for cognition. Cambridge: Cambridge University Press.

van Dijk, T. A., & Kintsch, W. (1983). Strategies of discourse comprehension. London: Academic Press.

Unconscious Inhibitory and Conscious Facilitatory Effects in Semantic Priming

F. Herbelleau & S. Delord

University of Lyon 2 (France)

delord@univ-lyon2.fr

Automatic/strategic semantic priming has been a major focus in the semantic priming literature, mainly because of the much debated question of perception without awareness. Namely, the point was to determine if the masked priming paradigm warrants a dissociation between access to stored semantic representations and consciousness. When the prime could be consciously identified, both automatic (i.e., spreading activation) and strategic mechanisms (i.e., prospective expectancy generation, retrospective processing) could account for the observed facilitatory and inhibitory priming effects. In order to avoid the strategic use of the prime by the subject and hence to isolate automatic mechanisms, a mask is classically added before and/or after a shorten prime. In this case, the mask is supposed to selectively disrupt the processing of the prime before the access to consciousness, without any perturbation of the lower steps of prime processing, such as the activation of stored representations. This conception is however difficult to reconcile with psychophysical models of visual masking based on low level interactions between the activities involved by the masked stimulus and by the mask. In the present work, visibility of the masked prime was manipulated by varying the prime duration while maintaining the delay between the prime and the target constant (1000 msec). This was done to test various levels of prime awareness, from subliminal perception (for the shorter durations) to conscious identification (for the longer ones), on long lasting priming. Prime awareness was assessed at each trial both by a detection task (primes were absent on one third of the trials) and by an identification task. The lexical decision on the target was used to evaluate the indirect priming by comparing three types of primes: A word semantically related to the target, a non related word and a no prime condition. The comparison between the two control conditions (i.e., non related prime and no prime) served to assess inhibition involved by the non related primes whereas the comparison between the related and the two control conditions served to assess inhibitory and facilitatory effects involved by the related primes.

Method

Two groups of 30 subjects participated in the experiment. The material comprised 75 target words (for instance, the French word COUTEAU) associated with a related word (lame), an unrelated word (rame), and a no prime condition (a blank) and 75 target pseudo-words (TIRPENT) with an unrelated word prime (fée) and a no prime condition. Fifteen experimental lists were created to counterbalance the type of prime (i.e., related, non related or no prime) and the five durations of the prime (6, 19, 31, 44, and 56 ms). These very short durations were due to the 160 Hz refresh rate allowed by the Matrox Millenium II graphic card and the Sony Multiscan screen we used. Each subject passed through three distinct experimental blocks (in a 1-2-3 or a 3-2-1 order, depending on the subject's group). A " masked prime + target " block comprised 150 trials with the following sequence at each trial: The fixation point, followed by the prime (for 6 msec up to 56 msec, from a trial to another) after which a pattern mask (XXXX) was presented for 1 sec minus the duration of the prime, and then the target appeared until the subject's responses: A lexical decision task, then both the detection and identification of the prime. In the "masked prime" block, the same trials were run in a different random order, except that the target was removed and that only the tasks pertaining to the prime were required. In the last "prime alone" block, the primes were presented

without a mask nor a target and, due to the increased visibility of the primes, shorter durations were used (6, 12, 19, 25, and 31 ms).

Results

The main result for the RT of lexical decision was an interaction between the type of prime and the duration of the prime. Whereas the classical facilitatory priming effect was observed for the longer prime durations, with shorter RT for the related prime/target pairs than for the non related pairs or than for the no prime condition, an inhibitory effect was found for the very short prime durations, with longer RT for the related condition relative to the two control conditions. No inhibition produced by the non related primes relative to the no prime condition was found whatever the prime duration. Analysis of the performance on the prime in the "masked prime + target" block showed that only the shortest duration (6 ms) could provide a subliminal condition of presentation if a lenient criterion is used: Around one third of the primes were detected and a very few of them could be entirely identified. As the duration of the prime increased, performance increased until detection (from the 44 ms primes) and identification (from 56 ms primes) were perfect. A retroactive priming effect was also found for the intermediate durations of the prime: The related primes were better detected and better identified than the non related primes. However, comparison of prime identification in the three experimental blocks showed that the prime identification, that was perfect whatever the duration in the "prime alone" block, decreased when the prime was masked, but that there was no benefit due to the presence of the target in the "masked prime + target" block, presumably because of a competition between the different tasks that had to be performed in this block.

Discussion

To conclude, these results showed first a dissociation between semantic activation and consciousness, as priming was also found for the condition of presentation of the prime that was subliminal. This implies that, rather than an early perturbation, masking should mainly consist in a high level attentional disruption. Second, the manipulation of prime duration independently of the prime/target delay allowed to distinguish between an unconscious inhibitory priming effect and a conscious facilitatory priming effect. The conscious facilitation is consistent with strategic mechanisms of priming, according to which, for instance, the subject should attend to a related target once he had consciously identified the prime. On the contrary, the unconscious inhibitory component of priming should reflect an automatic inhibition subsequent to a prior automatic activation that our long SOA prevented. To our knowledge, this result is the first observation for prime/target semantic relation of the inhibitory effect that has recently been found for graphical or phonological relation.

Verbs and Mental Lexicon: A Case System in Mind

R. Bussone & J-P. Rossi

University of Paris XI (France)

regisb@limsi.fr

Verbs are said to be meaningful when mediated in sentences or situations. In this paper, we explore the organization of verbs in the mental lexicon. Two experiments, an association table and a priming experiment with lexical decision task have been conducted with verbs as primes. Results suggest that verbs are structured around a case system in the mental lexicon. Contrary to the network of nouns, verbs do not seem to be hierarchically or semantically organized together. The structure of the mental lexicon is studied for a long time, but curiously, the organization of verbs in the lexicon isn't the subject of many researches. The focus is made by on-line comprehension, but rarely on the really semantic status of verbs in the lexicon. Yet, the status of lexical units is quite important to understand, at least, what is stored in mind or how verbs are organized without a sentential background. The two experiments reported here show how verbs are accessed in the mental lexicon. Previous data, particularly morphological priming, show that, in English and German, infinitive and irregular forms are stored and accessed directly. The experimental method generally used is repetition priming versus the experimental and control conditions. The fact is that it is quite difficult to perform what repetition priming means, if it is a lexical access, a semantic access, or merely a graphemic recognition.

Experiment 1

To insure our investigations to test what is really semantic and/or associative, Experiment 1 establishes a small association table. This method permits participants to give answers quite simply and naturally. 217 French verbs were randomly chosen and tested with 40 participants, which were asked to write down " the first word that came in mind " after reading the prime. The list of verbs was randomly presented. Data were gathered and analyzed by frequency and types of responses. A categorization (in terms of synonyms, antonym, or related for verb-verb associations, or kinds of arguments for the rest of the associated words) has been established to classify each response or group of responses. A representation of the associative network extracted from this association table was also analyzed and compared with one made from noun primes (Bussone & Rossi, 1997; Cornuéjols & Rossi, 1999). Several observations have been made. For example, associated verbs represent only a fourth of the total of responses. The major category is represented by the object complements. The more common verb-verb associations are synonyms, and antonyms are rare. Those empiric elements show a low verb-verb architecture, with few interconnections between them, but some central concepts which concentrate a great amount of responses (house, money or car for example); there are also some high response frequency verb-verb associations but quite isolated in the network. Conversely, the noun-noun architecture is far more organized, and interconnections are not centered on one element but three or four, which play a central role. For example, if we consider body parts, arm, leg, hand and foot are central elements.

Experiment 2

From this first analysis, we selected high associated words (from 6 to 27 out on 40, that is to say up to 15% of response). A second classification has been made, using two basic categories: Semantic associations and predicative associations. Selected data were used for a priming experiment with lexical decision task (Experiment 2).

Experimental material. 64 verbs have been tested as primes. There were 67 semantic and 61 predicative targets, which were distributed as indicated in Table 1.

Table 1: repartition of the types of targets for experiment 2

	Semantic	Predicative	All targets
Verbs	49 (73.13 %)	2 (3.28 %)	51 (39.84 %)
Nouns	15 (22.39 %)	57 (93.44 %)	72 (56.25 %)
Adj. / Adv	3 (4.48 %)	2 (3.28 %)	5 (3.90 %)
Total	67 (100 %)	61 (100 %)	128 (100 %)

Several criteria have been controlled: The repartition of frequencies between the two conditions were not significant (Chi″(44) = 46.196, NS); the association forces between nouns and verbs were equivalent. In order to test each target without seeing the same prime twice, four presentation orders have been set up. Each participant saw sixteen couples of items for the predicative condition, sixteen others for the semantic condition, thirty two couples for the control condition, and sixty four couples of verb-nonword of items for the nonword condition.

Procedure. Thirty two participants took part in this experiment, eight per presentation order. Each session was composed of a training set of eight couples of items. After this training, the experimental session began, preceded by an alert message. Participants were told to answer " yes " or " no " by pressing the " A " or " P " key (on an AZERTY keyboard); the " yes " answer was defined by the lateralized hand (" A " for left-handed, and " P " for right-handed), conversely for the answer " no ". Each trial was composed of an eye fixation item (a raw of ten stars) presented in the middle of the screen for 500 ms, followed by a 100 ms blank screen; then the prime appeared in the center of the screen for 170 ms, then a blank screen for 80 ms, then the target appeared until the participant gave a response (up to 3500 ms).

Results

Considering the two experimental conditions and the control condition, result indicate that there is a massive priming effect ($F(2, 62) = 10.68$, $p < .0002$). Considering the planned comparisons, the semantic versus control condition is significant ($F(1, 31) = 19.60$, $p < .0002$), as well as the predicative versus control condition ($F(1, 31) = 12.67$, $p < .002$). It is not the case when we compare the two experimental conditions ($F(1, 31) = 0.59$, NS). As our experimental factors were maybe not relevant, a second analysis has been made, using a different classification. We distinguished the nature of the target (verbs vs nouns and adjectives / adverbs), so that it was possible to establish two control conditions (one for the verbs, one for the nouns, representing 50% for each category). That was impossible in the first analysis because the relation between the prime and the target was crucial. In this case, we could only have a single control condition. Finally, we had a 2 (type of target: Noun or verb) x 2 (condition: Experimental or control) plan. Response times are indicated in table 2. The results indicated neither interaction ($F(2, 62) = 0.77$, NS), nor effect of the type of target ($F(1, 31) = 1.13$, NS), but a massive effect of the condition ($F(1, 31) = 20.08$, $p < .0001$).

Table 2: RT and SD for each condition.

	Experimental		Control	
	Verb	Noun	Verb	Noun

Response times	602.20	602.38	622.64	637.14
SD	87.86	99.08	102.60	88.64

Planned comparisons show a significant priming effect for verb-noun trials (F(1, 31) = 10.92, p < .003), as well as for verb-verb trials (F(1, 31) = 4.26, p < .05).

Conclusion

These experiments, as well as those conducted by other recent experiments (see Ferretti, T. R., McRae, K. & Hatherell, A., in press; McRae, K., Ferretti, T. R. & Amyote, L., 1997) confirm the fact that verbs are stored as thematic role concepts, and defined as situation schemas. Verbs seem to refer to prototypical situations, generalized and stored in memory as common knowledge extracted from what we experience in everyday life. Also, what our experiments confirm is that we also have a weak but true verb-verb organization. The fact that a priming effect occurs in both verb-verb and verb-noun conditions confirm that verbs stand a double status language: as lexical units, verbs are defined by one another (synonyms, antonyms and related verbs), as any lexical units (nouns are also defined by other nouns); as " thematic role concepts ", verbs have also a role of organizers between concepts and actions.

The Impact of Causal Markers on Expository Discourse Comprehension in L1 and L2

L. Degand* & T. Sanders**

*University catholic of Louvain (Belgium)

**Utrecht University (The Netherlands)

degand@exco.ucl.ac.be

In our presentation we want to report on an experiment dealing with the impact of linguistic markers, in particular causal connectives and signalling phrases, on the comprehension of expository texts. Although there has been a lot of work in this area, there is still no detailed consensus on the role of explicit discourse markers in text (for an overview, see Sanders & Noordman, 2000). In the literature it is assumed that they influence both the reading process and the reading result, i.e. the mental representation constructed by the reader. There is indeed much empirical support for the position that connectives and other linguistic coherence markers play a facilitating role during the reading process, i.e. on-line (e.g. Bestgen & Vonk, 1995; Sanders & Noordman, 2000). With respect to the influence of explicit coherence markers on the text representation afterwards, i.e. off-line, the positions are not so clear. On the one hand, there are studies which seem to show that linguistic marking of coherence relations improves the mental text representation (better recall, better comprehension, better verification) (e.g. Degand, Lefèvre, & Bestgen, 1999; Millis & Just, 1994). On the other hand, there are a number of studies indicating that linguistic markers do not influence the text representation after people have read the text (e.g. Sanders & Noordman, 2000; Spyridakis & Standal, 1987) or even have a negative impact

(Millis et al., 1993). There are plausible explanations for the reported contradictions, which can be brought back to two basic problems with most of this previous research: Sloppy manipulation of the independent variable and neglecting of other influencing factors like content knowledge and linguistic knowledge of the reader.

In our experiment we replicated the method followed by Degand et al., (1999) in which natural-like texts were used on topics that do not require any previous knowledge from the reader. The manipulation of such texts gave rise to a clear impact of linguistic marking, the presence of (causal) connectives leading to a better comprehension performance. In addition, we wanted to control and check the impact of the linguistic knowledge of the reader. To that end, we contrasted L1 and L2 readers, and we contrasted two different types of linguistic marking, i.e. causal connectives and causal signalling phrases. The latter was done in order to check whether the surface form of the linguistic marking could be of any influence on the comprehension process. In addition to the main hypothesis that texts with linguistic marking (connectives and signalling phrases) should lead to better comprehension than the implicit versions, we formulated three competing hypotheses concerning the impact of linguistic marking on L2 readers in contrast to L1 readers: (1) Both L1 and L2 readers benefit from linguistic marking, but L2 readers benefit more than L1 readers; (2) L1 and L2 readers both equally benefit from linguistic marking; (3) L2 readers do not or do hardly benefit from linguistic marking, while L1 readers do.

On the basis of the literature, all three hypotheses are plausible. Hypothesis 1 is based on the observation that L2 readers need more (explicit) help with text processing in a Foreign Language because they have trouble with higher-order processes of text comprehension. Hypothesis 2 refers to the so-called "inter-dependence hypothesis" (Cummins, 1984) that states that cognitive and linguistic skills acquired in one language can be straightforwardly transferred to another language. In our context this would mean that as soon as readers master an efficient reading strategy in their mother tongue, including the ability to utilise and infer logical relationships in discourse, they can transfer this skill to another language. Finally, hypothesis 3 is based on the assumption that a minimal level of linguistic proficiency is required to process (complex) discourse. L2 readers will have trouble understanding markers of the discourse structure, and for that reason they will not be influenced by the markers.

Method

In the experiment, participants (54) read expository texts of approximately 250 words and answered comprehension questions. The texts (18) were manipulated with respect to the presence or absence of linguistic markers: In some texts, we manipulated the presence or absence of causal connectives, in other texts, the presence or absence of causal signalling phrases. Doing this, we took care not to disturb the quality of the texts (adequately used and relevant connectives, texts that were long enough, etc.) so as to arrive at three versions of the texts that would be as natural as possible. The comprehension questions focused either on a manipulated relation (presence or absence of a linguistic marker) or on other parts of the text. The two languages under investigation were French and Dutch. Both languages functioned as L1 and L2 in the experiment, i.e. (French-speaking) Belgian students (group 1) read texts in French (L1) and in Dutch (L2), and Dutch students read the same texts in Dutch (L1) and French (L2).

Results

The results show an overall significant effect of the text condition. A contrast analysis between the three conditions furthermore shows that the implicit condition differs significantly from the explicit conditions, i.e. marking with connective or signalling phrase, while the explicit versions do not

significantly differ from each other. As a matter of fact, texts with explicit (causal) coherence markers lead to a better comprehension performance than texts without these markers. It is noteworthy that this effect persists throughout the nine L1 texts and the nine L2 texts (regardless of the language), i.e. they all showed an advantage for the cohesive version. In addition, there was no interaction between the text condition and the manipulation. It follows that the cohesive manipulation affected the questions about the marked relations, as well as the questions covering other parts of the text. When comparing the results for L1 and L2, we note a highly significant effect of the language, the scores for L1 being a lot better than those for L2. However, there was no interaction between L1 and L2 and the text condition. Given our results, this means that both in L1 and L2 the comprehension performance is affected by the text condition, more cohesive texts (explicit relational marking) giving rise to a better performance than less cohesive texts.

Discussion

We want to address two issues: First, we want to discuss the three hypotheses concerning the contrast between L1 and L2. While, given our results, hypothesis 1 can be excluded straightforwardly, hypotheses 2 and 3 are both worth exploring, since the detailed results of our experiment showed a slightly different scoring pattern for the two groups, group 2 being globally better and group 1 weaker for L2. Second, we want to address the question why we found an effect of linguistic marking while many other studies did not. Discussion of this issue includes theoretical as well as methodological considerations.

References

Bestgen, Y. & Vonk, W.(1995). The role of temporal segmentation markers in discourse processing. Discourse Processes, 19, 385-406.

Degand, L., Lefèvre, N. & Bestgen, Y. (1999). The impact of connectives and anaphoric expressions on expository discourse comprehension. Document Design, 1, 39-51.

Sanders, T.J.M, & Noordman, L.G.M (2000). The role of coherence relations and their linguistic markers in text processing. Discourse processes, 29, 37-60.

Text Cohesiveness and Feeling of Comprehension

N. Lefèvre & G. Lories

University Catholic of Louvain (Belgium)

lefevre@exco.ucl.ac.be

The literature on metacognition suggests that a large number of cues, possibly divergent, can support metacognitive judgments. While some are provided by the content of working memory after retrieval has been attempted, some are provided by what happens during processing. For instance Begg, Duft et al. (1989) suggest that meta-memory ratings are based on ease of

processing. A distinction between sources of information may be important also to understand metacomprehension. There is evidence that on-line, transitory information can be used to express ratings during reading. In a text comprehension task in which the subjects assess their comprehension segment by segment and ratings are collected on-line, Ehrlich and Cahour (1991) observe text cohesion effects on the comprehension ratings, at least on a second reading: The self evaluations (segment by segment) were higher for the cohesive version. This result suggests that difficulties encountered during reading, if they are remembered, may also contribute to decrease off-line comprehension judgment. In this case, ratings made after reading may rely in part on events that came up during reading. On the other hand, recent experiments (Degand, Lefèvre and Bestgen, 1999) show no impact of discourse markers on a final judgment of comprehension even though the markers did help make the necessary inferences during reading. This result goes the other way and suggests that on-line and off-line ratings should be dissociated. The present experiment investigates whether such a dissociation hypothesis is appropriate. One way the dissociation might happen is that non integrated elements may be present in working memory when the readers give their on-line judgment but the final judgment of comprehension may remain unaffected because problems have been solved by the time the readers came to the end of the text. This possibility actually exists within the construction-integration (CI) model of comprehension (Kintsch, 1988).

According to the CI model, the reader first builds an approximate representation of the text, that is moulded into a coherent representation only later on, through a specific integration process: "Filling the gaps" is thus considered usual at the end of the construction stage. This integration process is required even when the stimulus text is coherent because, for instance, some elements of information may always be missing in the world knowledge of the reader. The reader is then expected to keep thinking until the representation obtained appears coherent enough with respect to reading objectives. According to this view, satisfactory reading involves a comprehension judgment based on the quality of the final representation obtained, its coherence and its richness but this final judgment need not be affected by transitory aspects of processing. The transitory difficulties met when processing the text should be an ordinary fact of life and should not affect a final comprehension rating because, if they did, the judgment could have no adaptive value. Reciprocally, on-line assessments may be affected by processing problems that are not solved yet and leave conflicting elements in working memory.

According to this interpretation of the CI model we should observe a correlation between a final (off-line) feeling of comprehension and comprehension performance but the effect of processing difficulties should be mediated by the resolution they receive. The final comprehension judgment should be affected by reading difficulties that cannot be remedied and leave no possibility to build a coherent and rich representation. On the other hand, self evaluation by segment (on-line) may be affected by text cohesiveness because of unsolved elements in working memory.

We ran an experiment to assess the effect of text cohesion and self-paced vs constrained reading on comprehension and metacomprehension. The comprehension ratings were asked after each paragraph (on-line) and after reading the whole text (off-line). We expected the cohesion manipulation to affect only on-line ratings while the constraints on the pace of reading would affect both on-line and off-line ratings. The less cohesive texts and the constrained reading should make integration more difficult but cohesion problems can be solved during reading and thus should affect only on-line ratings.

Method

In the experiment, six expository texts of approximately 550 words divided in 4 paragraphs were taken from popular encyclopedias and adapted. They covered topics like Aztecs (Les Aztèques), optical fibers (Les fibres optiques), liver fluke (La douve du foie) etc. For each text, we prepared a cohesive version (control condition) and a less cohesive one (experimental condition) by

manipulating anaphora in two paragraphs (second and third paragraphs or second and fourth paragraphs). The cohesive texts use unambiguous anaphoric links in which the referential expression was a repetition of the antecedent. The less cohesive texts contained anaphors that were difficult to resolve. Fifty-six subjects read six texts each: Three in the control version and three in the difficult version. The subjects were warned that they would be asked comprehension and metacomprehension questions. There were two reading conditions: Self-paced or constrained. In the constrained condition the participants were allowed ninety percent of the mean reading time taken by a group of twenty subjects (for each segment) in a preliminary test. For each text, the participants answered eight comprehension questions (two questions about each paragraph): Two of them tapped the cohesion manipulation and the other six covered non-manipulated parts of texts (these were either detail or integration questions). The participants answered metacomprehension questions on-line and off-line (during and after reading) on a 0-20 scale: They gave a feeling of comprehension rating at the end of each paragraph and after reading the whole text.

Results and Discussion

Both manipulations (cohesion and pace of reading) affected comprehension in the expected direction: We observed a better comprehension performance for the cohesive version of texts and for the self paced condition. There was no interaction between these experimental manipulations for the comprehension measure. This result is imperfectly reflected in metacognitive ratings. Although metacognitive ratings correlated with comprehension performance (within subject) and although pace of reading affected all metacognitive ratings, cohesion manipulation affected only on-line ratings. This pattern of results supports the idea that participants evaluated their comprehension after reading, based on the representation built, when problems had been solved, and not on the difficulties encountered during reading. Finally, off-line ratings also showed no interaction between experimental manipulations (cohesion and pace of reading); Cohesion had no impact at all on the off-line rating. This suggests that subjects feel they have successfully solved cohesion problems.

References

Begg, I., Duft, S., Lalonde, P., Pelnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. Journal of Memory and Language, 28, 610-632.

Degand, L., Lefèvre, N., & Bestgen, Y. (1999). The impact of connectives and anaphoric expressions on expository discourse comprehension. Document Design, 1, 39-51.

Ehrlich, M.F., & Cahour, B. (1991). Contrôle métacognitif de la compréhension: cohésion d'un texte expositif et auto-évaluation de la compréhension. Bulletin de Psychologie, 44, 147-155.

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. Psychological Review, 95, 163-182.

Anaphoric Resolution and Discourse Focus : Referential Accessibility in question ...

M. Fossard*/, J-L. Nespoulous* & D. Cardebat****

*University of Toulouse-le Mirail (France)

** INSERM 455, CHU Purpan (France)

fossard@mail.univ-tlse2.fr

The concept of "Discourse Focus" (Sanford & Garrod, 1981), also called "Discourse Center " (Centering Theory, 1986, 1995, 1998), is based on the idea that the discourse entities mentioned or evoked in a discourse at any point in time, are not all in the same level, but some of them have a very particular prominence that appears, among other factors, with the resolution of certain anaphoric expressions. Following the Centering Theory of attention focus in discourse and basing our argument on the identifiying function of referential expressions, we formulate the assumption that if the anaphoric pronoun "il" is a signal to maintain attention on the focused entity (the high accessibility entity), a repeated noun would rather indicate a change of attention center towards a less focused discourse object (the low accessibility entity). This will be dealt with in the first experiment (experiment 1). Another experiment (experiment 2) was constructed with the aim of determining the referential functioning of the "discrete" demonstratif "celui-ci", insofar as this hybrid "anaphoric-deictic" expression (Kleiber, 1994) makes it possible - more specifically than a repeated noun - to change the focus of attention, while preserving local discourse coherence. In order to test these hypotheses, we compare target sentence reading time measures for each text condition.

Method

For each of the experiments, 40 young, right-handed subjects were tested individually. They had to read the 40 randomised experimental texts and 48 distractor texts on a computer screen. The texts consist of 3 sentences - the last being the target sentence - and they are followed by an assertion (true/false judgment) which relates to the target sentence except for the distractor. The texts present 2 characters one of whom is focused (entity 1), and the other one is not (entity 2). We used the following procedure: The subjects read the text in two successive stage. They had to push the space key to move from the first part of the text to the next. Then, they had to make a true / false judgment by using either the "F" or the "V" key.

Results

Tables 1 and 2 present the results of both experiments. We indicate the reading times (RT) taken to process the target sentences and the error rate by condition. The RTs have been corrected in order to take sentence length (number of characters) into account.

Table 1: Reading times (ms) and error rate (%) for each of the conditions of experiment 1

	pronoun "il"	repeated noun
Entity 1	38.5 (4.5 %)	40.6 (7.2 %)

Entity 2	44.8 (5.9 %)	41.9 (5.4 %)

Reading / comprehension times for the target sentences: The analysis of variance (realized by subject (F1) and by text (F2)) reveals a main effect of "entity type" on the RT (F1(1, 39) = 38.3, p < .0001; F2(1, 39) = 25.7, p < .0001), reflecting the fact that processing is faster when the target sentence refers to the entity 1 (M = 39 ms) rather than to the entity 2 (M = 43.4 ms). This result is in conformity with the important role of discourse focus on the processing of referential expressions: Refering to the focused entity (entity 1) clearly reduces the reading times. An anova shows an interaction of "entity type" and "type of referential expression" (F1(1, 39) = 8.8, p = .005; F2(1, 39) = 10.6, p = .002). Indeed, processing is faster for the focused entity (entity 1) with a pronoun than with a repeated noun (F1(1, 39) = 6.4, p = .01; F2(1, 39) = 4.9, p =. 03), and it is also faster for the pronoun with an entity 1-referent than an entity 2-referent (F1(1, 39) = 24.1, p < .0001 ; F2(1, 39) = 29.5, p < .0001). This result indicates that the pronoun is an indicator of discourse focus. On the other hand, processing is faster for the less focused entity (entity 2) with a repeated noun than with a pronoun (F1(1, 39) = 4.6, p = .03; F2(1, 39) = 7.8, p = .008), but there is no difference for the repeated noun with an entity 1-referent or an entity 2-referent. This result indicates that although the repeated noun is in a position to change the center of attention, it is not nevertheless, sensitive to entity focusing.

Table 2: Reading times (ms) and error rate (%) for each of the conditions of experiment 2

	pronoun "il"	pronoun "celui-ci"
Entity 1	49.6 (1,25 %)	60.5 (5.25 %)
Entity 2	61.7 (2,75 %)	49,8 (3.75 %)

Reading / comprehension times for the targets sentences: Unlike the previous experiment, there is no main effect of "entity type" on the RT in this experiment (F1(1, 39) = .374, p = .5446; F2(1, 39) = .098, p = .756): Refering to the focused entity (entity 1) does not improve the reading times with respect to the entity 2. However, there is an interaction of "entity type "and "type of referential expression" (F1(1, 39) = 41.1, p < .0001; F2(1, 39) = 122.8, p < .0001. These results indicate that the anaphoric-deictic "celui-ci" is sensitive to entity focusing (unlike the repeated noun in the first experiment).

General Discussion

The results reported here confirm our predictions. Although the anaphoric pronoun "il" seems to be a signal to maintain attention on the focused entity (entity 1), a repeated noun rather indicates a change of attention center toward a less focused discourse object (entity 2). However, the results show that the constraints regulating the referential functionning of the different expressions are not the same in each case. Indeed, repeated noun and pronouns are clearly in contrast: The interpretation of a repeated noun seems independant of its immediate context of use, which is not the case with indexical expressions like "il" or "celui-ci" which are marked as having active referents. In other words, the anaphoric pronoun "il" appears to act like a signal of "high focus",

whereas the anaphoric-deictic "celui-ci" acts like a signal of "low focus". Repeated nouns appear to be neutral in relation to entity focusing. In other respects, these results seems to indicate that the notion of discourse focus is apparent only after the re-identification of a given referent (indeed, the results of experiment 2 do not show a "type of entity" effect - the entity 1, per se, is not more accessible than the entity 2 -). This could be indicate that the notion of discourse focus is not independant of the different referential expressions, but rather autonomous. In sum, the discourse structure could give an entity preference (entity 1), but this preference must be confirmed by the appropriate referential expression (the anaphoric pronoun). To change rapidly the focus of attention (i.e., without additional cost), it's necessary to use a specific expression (like the anaphorico-deictic pronoun).

References

Cornish, F. (1995). Références anaphoriques, références déictiques, et contexte prédicatif et énonciatif. In Sémiotiques, "anaphores : marqueurs et interprétations", dirigé par Kleiber, G., CNRS, 8, 31-57.

Garrod, S., Freudenthal, D., Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of sentences in a discourse. Journal of memory and language, 33, 39-68.

Gordon, P., Grosz, B., Gilliom, L. (1993). Pronouns, nouns and Centering in discourse. Cognitive Science, 17, 311-347.

Grosz, B., Joshi, A., & Weinstein, S. (1995). Centering : A framework for modeling the locale coherence of discourse. Computational Linguistics, 21/2, 203-225.

Kleiberg, G. (1994). Anaphores et pronons. Champs linguistiques. Duculot.

Sanford, A., & Garrod, S. (1981). Understanding written language. Wiley Edition

Walker, M. (1998). Centering, anaphora resolution and discourse structure. In Centering in Discourse. Walker, Joshi & Prince (eds). OUP. 1-29.

Idiom Processing : Effect of Compositionality

S. Caillies *, K. Butcher * & W. Kintsch *

***University of Reims (France)**

****University of Colorado at Boulder (U.S.A.)**

stephanie.caillies@univ-reims.fr

It's commonly assumed that idioms have lost their metaphoricity over time and now exist as frozen metaphors. Indeed, an idiom has been traditionally defined as a locution for which the intended meaning is not derived from the meaning of the individual words comprising it (Swinney & Cutler, 1979). In this framework, the figurative meaning of the idiom "Armed to the teeth" can not be

derived from the compositional analysis of "Armed" and "Teeth". Because the comprehension of idioms can not be explained by standard models that assume meaning is derived from a strict compositional analysis of the linguistic input, special models of idiom comprehension have been proposed, noncompositional and compositional models. Noncompositional models assume that idiomatic expression is identified like a lexical entity stored in memory. Although noncompositional models which have been proposed differ in terms of how and when meaning of idioms is retrieved, they share the assumption that idiomatic meaning is semantically distinct from the literal meanings of the component words. Contrary to noncompositional models, compositional models assume that idiom meaning is not encoded as a separate lexical entry but as a meaning that is associated, either with a particular configuration of words (Cacciari & Tabossi, 1988) or with pre-existing conceptual metaphors (Gibbs, Boddanovich, Skyes, & Barr, 1997). In this framework, when an idiom is encountered, the expression is processed literally, and at the same time may emerge an idiomatic meaning. These two compositional hypotheses, configurational and metaphorical, share the assumption that some relationships exist between idiomatic meaning and idiom parts meaning, and consequently that the frozen metaphor view of idiomaticity is wrong (Gibbs, 1992). Consistent with compositional models, a growing body of literature supports the assumption that idioms can differ in the degree to which literal word meanings contribute to an idiomatic meaning (Gibbs, Nayak, & Cutting, 1989; Titone & Connine, 1994 a), and that this degree of compositionality exerts not only an effect on idiom comprehension (Gibbs, 1991; Levorato & Cacciari, 1999) but also on idiom processing time (Gibbs, Nayak & Cutting, 1989). These researches have demonstrated that a strict semantic dichotomy between literal and figurative meanings is inadequate to account for idiom processing, and that the frozen metaphor view is wrong, at least for the decomposable idioms in which a part of the figurative meaning is used literally. The main purpose of the present study was to further examine the psychological relevance of the compositionality factor. Our objective was to investigate its effect on idiom processing, and more precisely its effect on the time course of activation of literal and figurative meanings of ambiguous idioms. We predicted that the figurative meaning would be more easily and more quickly activated during the reading of decomposable idioms than during the reading of nondecomposable ones. Indeed, we assumed that a semantic relationship between the meaning of the individual words comprising the idiomatic expression and the figurative meaning would facilitate the processing of the idiom and the retrieval of the figurative meaning. Moreover, as it has been shown in several studies (Cacciari & Tabossi, 1988; Titone & Connine, 1994 b), we expected that the idiom always would be processed literally.

Method

In our study, participants were instructed to read decomposable and nondecomposable idioms in a fixed time, and to perform a lexical decision task either 50 ms or 350 ms after reading. These idioms were preceded by a neutral context, biasing neither the idiomatic nor the literal interpretation of the idiomatic phrase. They were chosen from the descriptive norms published by Titone and Connine (1994a), and were highly familiar and highly literal; that is, there was high potential for the idiom to receive a literal interpretation. Three types of targets were used: Control, literal and figurative. The idiomatic and literal targets were selected by analyzing participants' responses to association tasks. The control targets were selected in such a way they were neither related to idiomatic nor to literal meanings of the experimental idioms. Moreover, half of them matched on frequency, length, and syntactic structure to figurative targets, the remainder matching on the same dimensions to literal targets. These experimental idioms were mixed with fillers, and for 35% of them, a comprehension question was presented subsequent to the lexical decision task to make sure that participants understood the sentences they were reading.

Results

The results mainly showed an interaction between compositionality and target factors ($F(2, 70) = 4.38$, $p < .02$) but no significant effect of delay. This interaction indicated that (a) the difference between the lexical decision times of the figurative targets and those of the control targets was positive for the decomposable idioms (78 ms) and negative for the nondecomposable ones (-113 ms); (b) the difference between the lexical decision times of the literal targets and those of the control targets was negative and smaller for the decomposable idioms (-35 ms) than for the nondecomposable ones (-113 ms). Consistent with our prediction, the results suggest that during the processing of this kind of idioms, the literal meaning is processed, and that the compositionality factor has an effect on the processing of the figurative meaning. However, contrary to our expectation, the figurative meaning is activated in this experiment only when the idiom read is a decomposable ones.

Conclusion

These results allow us to conclude that the frozen metaphor view of idiomaticity is wrong since the semantic relationship between the literal meanings of the component words of the idiom and the figurative meaning has an effect on the time course of activation of figurative meaning: It does not facilitate its activation. Indeed, the figurative meaning of decomposable idiom needs more than 350 ms to be activated whereas that of nondecomposable ones is activated almost as soon as the idiom is processed.

References

Cacciari, C., & Tabossi, P. (1988). The comprehension of idioms. Journal of Memory and Language, 27, 668-683.

Gibbs, R. W. (1991). Semantic analyzability in children's understanding of idioms. Journal of Speech and Hearing Research, 34, 613-620.

Gibbs, R. W. (1992). What do idioms really mean? Journal of Memory and Language, 31, 485-506.

Gibbs, R. W., Nayak, N. P., & Cutting, C. (1989). How to kick the bucket and not decompose: Analyzability and idiom processing. Journal of Memory and Language, 28, 576-593.

Gibbs, R. W., Bogdanovich, J. M., Sykes, J. R., & Barr, D. J. (1997). Metaphor in idiom comprehension. Journal of Memory and Language, 37, 141-154.

Levatoro, M. C., & Cacciari, C. (1999). Idiom comprehension in children : Are the effects of semantic analyzability and context separable? European Journal of Cognitive Psychology, 11, 51-66.

Swinney, D. A., & Cutler, A. (1979). The access and processing of idiomatic expressions. Journal of Verbal Learning and Verbal Behavior, 18, 523-534.

Titone, D. A., & Connine, C. M. (1994a). Descriptive norms for 171 idiomatic expressions: Familiarity, Compositionality, Predictability, and Literality. Metaphor and Symbolic Activity, 9, 247-270.

Titone, D. A., & Connine, C. M. (1994b). Comprehension of idiomatic expressions : effects of predictability and literality. Journal of Experimental Psychology: Learning, Memory and Cognition, 20, 1126-1138.

Why Should We Use Connectives in Discourse?

C. Soria

Institute for Computational Linguistics – CNR (Italy)

soria@ilc.pi.cnr.it

In this paper we are concerned with some aspects of the use of connectives in discourse. Connectives are widely considered as explicit lexical realisations of semantic/pragmatic relations holding between adjacent discourse segments, at any level of granularity (Hobbs, 1985, 1990; Knott & Mellish, 1996; Sanders et al., 1993). The role of connectives is often described as instructional in essence, as a clue for speakers and hearers during the process of establishing a coherence relation between two discourse segments (Noordman & Vonk, 1996; Caron, 1987). At the same time, it is acknowledged that the occurrence of a connective is not a prerequisite for a person to infer the coherence relation holding between the discourse segments: In other words, connectives would be optional, their use being a matter of individual preference, topic and style (Schiffrin, 1987). The main focus of our study was to contend this assumption, by giving experimental evidence of several possible sources of constraint on the way in which the frequency connectives are used in discourse. The two hypotheses that lie behind our work are as follows: First, occurrence vs. non-occurrence of a connective could be related to the particular type of coherence relation which is to be inferred by any receiver of a given message; Second, occurrence vs. non-occurrence of a connective could also be due to the particular constraints which are imposed by the particular way of expression (spoken or written) that is chosen.

Coherence relations have been addressed by a number of different approaches for the purposes of text understanding, text analysis and text planning. This results not only in a difference in the theoretical interpretations of the concept, but also in a heterogeneity among the actual sets of relations which are assumed to be psychologically and/or descriptively adequate. On the other hand, the various approaches show a considerable overlap in the individual relations adopted, thus suggesting a "common ground" shared by all approaches. In our view, coherence relations are to be conceived as cognitive operations that are activated by any interpreter of a message whenever there is the need to make two discourse segments connected. Our taxonomy of coherence relations consists of three broad classes or types of relations, each identified on the basis of the basic cognitive operation which is implied in the instantiation of a connection between adjacent segments of a discourse. We thus distinguish among additive, consequential and contrastive relations. Additive relations are based on some kind of similarity between the events, properties and relations represented by two sentences S1 and S2. Typical instances of additive relations are similarity, elaboration, list, exemplification and generalization relations, typically signalled by connectives such as and, or, or else, in other words, that is, or better, I mean, in addition, also, alternatively, for example, etc. Consequential relations involve establishing an implication or dependence (be it temporal, causal, modal, etc.) between the events, properties and relations represented by the two sentences. Typical instances of this type are causal, temporal or if-then relations, typically marked by connectives such as because, so, thus, then, after, if..then, etc. Relations belonging to the contrastive type, such as contrast, violated expectation or concession,

share the common feature of implying a negation of either similarity or implication between the two events, thus representing the negative counterpart of additive and consequential classes. Their typical markers are but, instead, on the other hand, however, etc. Involving a negative component, we hypothesize that contrastive relations are cognitively more complex than the additive and consequential types.

Experiment 1

Method

Our first experiment was addressed to inquiry the extent to which occurrence vs. non-occurrence of connectives could affect comprehension. In this experiment ten subjects were asked to give a spontaneous label expressing the particular coherence relations they perceived between two discourse segments, for a set of 72 examples, each containing a connective. Examples were taken from real texts of various genres. Each subject was then presented with the entire set of discourse segments' pairs, where a half contained the original connective, which had been removed in the other half. Cross-checking was ensured. It was then measured the extent to which subjects were able to correctly recover the coherence relation holding between two segments, and responses in those cases where a connective occurred were compared to responses to the same contexts, but with the connectives removed.

Results and Discussion

The data show that, in general, comprehension is significantly reduced if connectives do not occur. Any type of coherence relation is better understood when it is indicated by a corresponding connective. However, reduction in comprehension shows distinct patterns according to the particular type of relation involved: Comprehension is maximally affected in the case of coherence relations of the contrastive type, while it is only slightly affected for relations of the additive type.

The difficulty with which a contrastive relation is inferred without a connective could be due to the fact that explicit lexical marking is essential for the inference of the contrastive component of meaning to be made. According to our hypothesis, the contrastive type should be cognitively more complex than the other two, since it involves an additional negative component. If there is lack of marking, the negative or contrastive component of the relation is lost and the relation is thus interpreted as an additive or consequential relation, according to the underlying type. This is shown by the redistribution pattern into these other two classes. Lexical marking seems to have a rather shallow function for guiding the inference of additive relations: In fact, there is only a slight difference between the number of correct answers in the condition with connective and those in the condition without connective. This could be due to two different, non competing factors: First, it could be that additive connectives are not particularly efficient for guiding the inference of the relation because of their intrinsic ambiguity. Second, it could also be the case that interpreters tend to instantiate whenever possible a ''stronger'' implicational relation between two segments. Numerous studies (see Townsend & Bever 1978, Trabasso and Sperry 1985, Trabasso & Broek 1987, Garnahm et al. 1996, Noordman & Vonk 1997) suggest that the various types of relations differ not only along the dimension of cognitive complexity, but also along the dimension of cognitive relevance. According to these studies, causal and temporal relations are cognitively more relevant since they promote the storage in memory of the connection between information. On the other hand, the ease with which an additive relation is inferred when there is no lexical marking to signal it could be explained by the fact that a relation of this type is inferred by default, since it signals a simple and general continuation without any other particular connotation.

Consequential relations, while showing a reduction in the number of correct answers from the condition with connective to the condition without connective, display a relative ease of comprehension. This is not plausibly explained by making reference to the basic common knowledge shared by writers and readers, since this should determine a good recognition of contrastive relations as well. It is more advantageous to hypothesize that the high degree of understanding of consequential relations is related to the well-known preference toward inferring an implicational relation between the segments (see e.g. Black & Bern 1981). Again, the fact that this class of relations is the best comprehended one with lexical signalling is explainable in terms of semantic non-ambiguity of the connectives usually employed for its expression.

Experiment 2

The second experiment was aimed to verify the results obtained with the previous experiment, by inquiring the particular patterns of use of connectives in discourse production. For this experiment, 19 subjects were asked to tell a story depicted in a set of cartoons. Subjects first told the story in speaking, and then re-phrased it in writing. Results show that connectives are globally more frequent in speaking than they are in writing (M = 10.6% vs 7.6%). However, the respective patterns in which they occur according to the different types again indicates that the particular type of coherence relation constrains the occurrence of a connective. Additive relations are marked in around 20 to 25% of cases, consequential relations in 43 to 60% of instances, and contrastive relations with a frequency near to 100% in both versions of the narratives.

General Discussion

Results of both experiments thus indicate that connectives not only constrain the inference of coherence relations between two segments, but they are in turn constrained by the type of the underlying coherence relations and by the way in which a discourse is delivered (oral vs. written). It is particularly significant that connectives are almost obligatory when a constrastive relation is involved. We explain this result in the light of a higher cognitive complexity of contrastive relations. The much less obligatory use of connectives for additive and consequential relations can be similarly explained along the lines of cognitive relevance and complexity.

Overspecification of Referential Expressions in Instructive Discourse

A. Arts*, A. Maes*, L. G. M. Noordman* & C. Jansen**

*Tilburg University (The Netherlands)

**University of Nijmegen (The Netherlands)

Anja.Arts@kub.nl

This study focuses on the linguistic form of referential expressions in written instructive discourse. In making predictions for these linguistic forms, recent theories on discourse reference generally depart from two basic concepts. Central in the first concept is the idea that the form of a referential expression is indicative of the degree of mental activation (accessibility or givenness) of

the entity or referent the expression refers to (Ariel, 1991; Gundel, Hedberg & Zacharski, 1993). The second basic concept is referential economy: Producers of language do not usually provide more information in a referential expression than what is necessary for unique identification of the referent (Grice, 1975).

Mental activation and referential economy are important determinants of the linguistic form of referential expressions in discourse, both when the expression is the initial reference to an entity and when the expression is a subsequent reference (an anaphor). However, under the influence of different textual circumstances and language-in-use factors, referential phenomena in discourse can deviate from the pattern predicted by these basic concepts. This contribution discusses one such deviation: The overspecification of referential expressions in instructive texts. Overspecified referential expressions are expressions that contain more information than what is necessary for unique identification of the text referent. The referential phenomena that are presented originate from a production experiment. Subjects were given the task to write an instruction for a radio alarm. To enable them to do this, the experimental material contained very precise information about the different functions of the radio alarm. The larger part of this information was not given in verbal form, in order to rule out a possible effect on the subjects, formulation of the instructive text. As a result, all action information was given visually and only the goal information was given verbally, in minimal wording (for instance " set hour ").

As an independent variable, the assumed user's condition was manipulated in the experimental task. One group of subjects (n = 27) was asked to write the instruction from the assumption that the intended reader was to set the radio alarm only once, that same evening (reading to do). A second group of subjects (n = 21) was asked to write the instruction from the assumption that the intended reader was to set the radio alarm frequently, every evening, and therefore needed to learn the task of setting the alarm (reading to learn). It was the expectation that the proportion of overspecified referential expressions would be higher in the " reading to learn " condition, because the subjects in this condition were expected to anticipate their reader's need for additional (declarative) information. We analysed the linguistic forms of the initial referential expressions and the anaphoric expressions in the experimental corpus. The analysis showed that the linguistic forms in the corpus do not mirror the predictions for these forms that result from the concepts of mental activation and referential economy.

The initial referential expressions were often overspecified. We analysed 311 initial expressions that occurred in action sentences (sentences in which a particular action was explained, that was to be executed with one of the objects on the radio alarm). In every action sentence one of the objects on the radio alarm (the button to push, or the switch to turn) needed to be referred to in an expression that enabled the reader to uniquely identify the object. The analysis showed that the writers often provided more information than what was necessary for unique identification of the object; 57 per cent of the initial expressions was overspecified.

In the same vein, the anaphoric expressions were often overspecified. The design of the experimental material elicited anaphor usage in many segments of the texts that were written. However, in two segments the experimental material elicited direct anaphor usage. By this we mean that the initial reference to an object was directly followed by the anaphoric reference to the same object. This occurred when, in successive action sentences, two different actions needed to be explained, for which the same object, and only that one object, was to be used. These text segments resulted mostly in two action sentences within one paragraph, the first action sentence containing the initial reference to the object (the antecedent), and the second action sentence containing the anaphoric expression. In (1) an example paragraph is given.

(1) If the alarm sounds, you press " the second button " once. If you want to turn off the alarm indefinitely, then press " the second button " twice.

In (1) the writer did not need to provide a repetition of the antecedent to ensure a correct link between the anaphor and its antecedent. A lesser specific anaphor (for instance " that button ")

would have sufficed as well. An analysis of 76 antecedent-anaphor sets showed that the anaphors often contained more lexical information than what was necessary to correctly link the anaphor to its antecedent. The writers regularly overspecified the anaphoric expressions; 42 per cent of the anaphors was overspecified. Anaphoric distance (the number of finite clauses between the antecedent clause and the anaphor clause) proved not a determining factor.

Independent variable. The independent variable had an effect on the production of overspecified referential expressions. The writers in the " reading to learn " condition produced more overspecified expressions than the writers in the " reading to do " condition. The differences were statistically significant, both for the initial referential expressions and for the anaphoric expressions. The analysis of the initial expressions showed that in the " reading to learn " condition 69 per cent of the expressions was overspecified. The corresponding figure in the " reading to do " condition was 48 per cent. As for the anaphoric expressions, 58 per cent was overspecified in the " reading to learn " condition. The corresponding figure in the " reading to do " condition was 30 per cent. We relate the results of the analysis to the principle of " distant responsibility " that was introduced by Clark and Wilkes-Gibbs (1986) and that is claimed to be applicable in non-feedback situations, when speakers or writers are distant from their addressee. This principle is derived from the principle of " mutual responsibility " which Clark and Wilkes-Gibbs introduce as well, and which is applicable in feedback situations. The point of departure in principle of " mutual responsibility " is that the correct resolution of a referential expression results from the collaborative effort of both dialogue partners, and is thus a shared responsibility. Clark and Wilkes-Gibbs, results show that a dialogue partner sometimes even provides an underspecified (insufficient) referential expression because he knows that if identification of the text referent proves to be too difficult, his dialogue partner will ask for clarification.

In the production experiment the writers were distant from their addressee; the principle of " distant responsibility " applies to these texts. Because requests for clarification are not possible, the collaborative effort is not equally divided between the discourse participants; most effort will be expended on the part of the writer. Also, the correct resolution of referential expressions is of the utmost importance in these texts. Failure to identify the objects on the radio alarm will lead to failure to execute the task. As a result the writers feel very responsible toward their reader and are more than willing to take upon themselves all of the (collaborative) effort that needs to be expended for successful completion of the referential process. This often results in overspecified expressions. The foregoing is applicable in both conditions of the experiment, but possibly more so in the " reading to learn " condition. In that condition the texts are written under the assumption that the intended reader will set the radio alarm every evening. Naturally, the consequences of an inadequate instructive text are most severe in this condition. In conclusion, this study shows that the concepts of mental activation and referential economy, which generally form the basis for predictions about the linguistic form of referential expressions, are subject to pragmatic constraints that characterise certain communicative settings.

Causal Factors that Influence Story Agents, Blame and Responsibility

B.A. Olde & A. C. Graesser

The University of Memphis (U.S.A.)

baolde@memphis.edu

In the field of discourse psychology, researchers have investigated how readers construct causes of action and events in narrative text (Graesser, Singer, & Trabasso, 1994). When a person mentally constructs a representation of actions, events, and objects, that person is building a situation model. Research in discourse psychology shows that time, space, protagonist, causality, and

intentionality are all important independent dimensions used in the construction of situation models (Zwaan, Langston, & Graesser, 1995; Zwaan, Magliano, & Graesser, 1995).

In the field of social psychology, theories of attribution investigate the process of explaining causes of human action (Hilton, 1990; Hilton, Smith, & Kim, 1995). Any given action may be caused by traits of the actor, by features of the situation, or by a complex interaction between traits and situations. For example, suppose that a father slaps his child's hand. Did he do this because he is normally a violent person (a pervasive trait), because he has the immediate goal of disciplining the child, or because he is desperately trying to prevent the child from grasping a loaded gun (a situation)? Attribution is the process of ascribing causes to human actions.

While research in attribution and causation have looked at many of the same variables that have interested discourse psychologists, the two fields sometimes reach different conclusions about which variables are more influential. Furthermore, it is hard to verify some of the assertions made because there are few studies that look at these factors together in any one experiment. Therefore, this study was conducted to clear up disagreement over the influence of causal distance, motive, intention, and outcome on decisions of causality, blame, and retribution.

This study used a 2 x 2 x 2 x 2 factorial design. The independent variables were motive, intention, outcome, and causal proximity. The dependent variables were ratings of causality, blame, and retribution. Participants read a packet containing 16 different stories. Each story included three agents: The primary agent, the main character and most active agent in the story; The intermediary agent, who separates the primary agent from the receiving agent; And the receiving agent, the agent that the eventual outcome happens to.

In attribution and causation research, researchers disagree about which agent in a causal chain of events is more responsible for the eventual outcome. It has been argued that the agent who is mechanistically closest to the outcome is more responsible (Einhorn & Hogarth, 1986; Johnson & Drobny, 1985; N'gbala & Branscombe, 1995; Shaver, 1985). However, it has also been asserted that the instigating agent is a greater influence than later agents, because the initial agent's actions may determine all of the subsequent actions (Girotto et al., 1991; Johnson et al., 1989; Vinokur & Ajzen, 1982; Wells et al., 1987). In this study, distance was not a factor for ratings of the primary agent; It was a factor for ratings of the intermediary agent; And it was only a factor for blame ratings of the receiving agent. The variability of these results are consistent with previous research that suggest people do not rely on any simple rule of causality (Graesser & Hemphill, 1991; Graesser, Lang, & Roberts, 1991; Spellman, 1997).

Studies have found evidence that the outcome of an event is an important influence when ascribing blame (Alicke, Davis, & Pezzo, 1994; Allison, Mackie, & Messick, 1996; Burger, 1981; Spellman, 1997; Walster, 1966). In this study, outcome had an effect on all ratings across all agents. These findings indicate that participants rated all agents involved more blameworthy simply because the outcome was negative.

Other researchers have focused on motive and intention as important determiners for making attributions of blame (Finkel & Groscup, 1997; Robinson & Darley, 1995). In this study, intention was also an important factor. If the primary agent's intention was negative, then ratings of the primary agent were higher for cause, blame, and retribution. However, researchers disagree about whether the event's outcome is more important than the active agent's intent or motive. In this study, when the primary agent's intention was negative and the story outcome was negative, there were significantly greater cause and blame ratings for the primary agent. Thus, as previous research has shown, both of these factors are important. But when intent and outcome were pitted against each other in a planned comparison, stories with negative outcomes produced significantly higher retribution ratings than stories with both a motive and intent. Thus, outcome had greater influence than the combined influence of motive and intent.

In summary, proximity can have an affect, depending on which agent participants rated; Cause, blame, and punishment ratings were all higher when the agent's intentions were negative; And all of the agents in the negative outcome stories received higher cause, blame, and punishment ratings.

References

Alicke, M. D., Davis, T. L., & Pezzo, M. V. (1994). A posteriori adjustment of a priori decision criteria. Social Cognition, 12, 28-308.

Allison, S. T., Mackie, D. M., & Messick, D. M. (1996). Outcome biases in social perception: Implications for dispositional inference, attitude change, stereotyping, and social behavior. In M. Zanna (Ed.), Advances in Experimental Social Psychology (Vol. 28) San Diego: Academic Press.

Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. Psychological Bulletin, 99, 3-19.

Finkel, N. J., & Groscup, J. L. (1997). When mistakes happen: Commonsense rules of culpability. Psychology, Public Policy, and Law, 3, 65-125

Girotto, V., Legrenzi, P., & Rizzo, A. (1991). Event controllability in counterfactual thinking. Acta Psychologia, 78, 111-133.

Graesser, A. C., & Hemphill, D. (1991). Question answering in the context of scientific mechanisms. Journal of Memory and Language, 30, 186-209.

Graesser, A. C., Lang, K. L., & Roberts, R. M., (1991). Question answering in the context of stories. Journal of Experimental Psychology: General, 120, 254-277.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comphrehension. Psychological Review, 101, 371-395.

Hilton, D. J. (1990). Conversational processes and causal explanation. Psychological Bulletin, 107, 65-81.

Hilton, D. J., Smith, R. H., & Kim, S. H. (1995). Processes of causal explanation and dispositional attribution. Journal of Personality and Social Psychology, 68, 377-387.

N'Gbala, A., & Branscombe, N. R. (1995). Mental simulation and causal attribution: when simulating an event does not affect fault assignment. Journal of Experimental Social Psychology, 31, 139-162.

Robinson, P. H., & Darley, J. M. (1995). Justice, liability, & blame: Community views and the criminal law. Boulder, CO: Westview.

Shaver, K. G. (1985). The attribution of blame: Causality, responsibility, and blameworthiness. New York: Springer - Verlag.

Spellman, B. A. (1997). Crediting causality. Journal of Experimental Psychology: General, 126, 323-348

Wells, G. L., Taylor, B. R., & Turtle, J. W. (1987). The undoing of scenarios. Journal of Personality and Social Psychology, 53, 421-430.

Zwaan, R. A., Langston, M. C., & Graesser, A. C. (1995). The construction of situation models in narrative comprehension: An event-indexing model. Psychological Science, 6, 292-297.

Zwaan, R. A., Magliano, J. P., & Graesser, A. C. (1995). Dimensions of situation model construction in narrative comprehension. Journal of Experimental Psychology: Learning, Memory, and Cognition, 21, 386-397.

Measuring Causal Force in Causal Chains

A. Guha & J-P. Rossi

University of Paris XI (France)

guha@limsi.fr

Causality is a fundamental dimension of coherence of mental representation of text, together with time, protagonists, goals, and space. It has however a special status, since a causal relation also implies particular temporal and spatial relations. The aim of our work was to explore the dimension of causality, through the notion of causal force of a cause-consequence couple. Keenan, Bailet and Brown (1984) showed that for reading couples of sentences, the reading time of the second sentence increased as the causal force of the couple decreased (causal force is estimated by likeliness judgments). This indicates that integration of the consequence is more costly if the causal force is weaker. van den Broek (1990) related causal force to logical conditions, such as necessity and sufficiency. We studied the relation between causal force and causal chains, in order to understand how causal force may vary in relation to the degree of detail (the number of intermediate steps), the logical relation between cause and consequence being constant. The main question we try to answer is the following. What are the features of causal force with regard to composition of causal relations? Consider, for example, that A is a cause of B (we note A -> B), and B -> C. Then, defining causality as "necessity under the circumstances", it is true that A -> C. Causal couples may logically compose, but is the same true of causal force? If we consider probability, it is true that the probability of (A,C) is the product of the respective probabilities of (A,B), and (B,C) (and hence is smaller than each individual probability). But this fact which is true in the world may not apply to the mental organization of causal strength. One could expect, for example, that in an explicative text which describes a causal mechanism (say, A->B->C->D->E->F->G), there may be a "macrostructural causal path" (A-D-G), and a "detailed causal path"(A-B-C-D-E-F-G). Then one could expect that under certain circumstances, a causal couple of the "macrostructural path" (for example (A,D)) could have the same force as (or a greater force than) one of its components in the detailed path (for example (A,B) or (B,C)). This would contradict the following hypothesis that can be formulated about causal force, analogous with the composition rule of probability: "for new information, the less detailed (or decomposed) the causal relation, the weaker the mental causal force".

Method

In order to test the validity of this rule, we collected likeliness judgments for cause-consequence couples of various detail levels, all taken from explicative texts. We collected 6 explicative texts describing natural phenomena in the style of scientific vulgarization, and extracted from each text a 7-sentence causal chain, with no linguistic marks of causal relation. Six versions of each text were elaborated in which various combinations of intermediate steps were skipped, hence varying the detail level of causal couples of the texts. For example, one version consisted of the seven sentences A-B-C-D-E-F-G (6 cause-consequence couples), and another consisted of just A-D-G (2 couples, namely A,D and D,G). We asked 86 subjects to judge likeliness for each couple of the six texts presented to him (each text in a different version). For example, if one text was presented to a given subject in its A-D-G version, the subject judged for this text the likeliness of A-D, before that of D-G (knowing A). Judgments of couples of a given text were mixed with those of the other texts, in a random order. The likeliness judgments of 15 couples for each text were collected (likeliness judgments ranged on a scale from 0 to 6), so the total number of couples whose force was estimated is 90. Each couple is characterized by its size, its original position (the position of the consequence in the original text), and by the actual position of the consequence in the version of the text presented to the subject. The size of a couple corresponds to the number of steps for decomposing it in the original text. In the earlier example, the size of A-D is 3, the original position 4, and the actual position is 3.

Results

Analysis of the data collected shows a significant effect of size on causal force, which decreases as size grows. It confirms - for these explicative texts - the hypothesis that the force of the couple resulting from the composition of two couples is weaker than that of each individual couple. However, contrast analysis doesn't show a significant difference between sizes 1 and 2 for causal force. More specifically, text by text analysis, or causal chain analysis doesn't show either the non-monotonous relation between causal force and causal chains that could be expected if the hypothesis of a "macrostructural causal path" was true. The original position and actual position of the consequence (which are correlated) also have a significant effect on force, but force decreases as original position is higher, whereas it shows a slight increase with actual position. This shows that if the effect of size resulted from a position effect, it could be only due to the fact that couples whose consequence has a high original position (towards the end of the original text) are by themselves less likely, and more frequent in large-size couples. In order to address this question, we analyzed the effect of size at constant original position, i.e. various couples whose consequences all have the same rank in the original text. Again, the effect of size is significant in all cases, and tends to prove that causal force is greater for "small" (more detailed) couples than for "large" (less detailed) ones. In order to find all determinants of causal force, we also encoded, for each couple, 15 variables, like the level of sufficiency, as well as temporal or spatial discontinuities, but also other aspects of relations between cause and consequence, like temporal saliency, temporal scale, or reversibility of the consequence. For all couples, the status regarding necessity was equivalent. Then we carried out multiple correlation analysis, which showed that the variables with significant impact (significant beta) on causal force, excepting the size of the couple, were the sufficiency level, spatial discontinuity, and reversibility, and the naturality or artificiality of the objects implied by the causal relation. Among these variables, sufficiency has the lowest tolerance (.30), which means that it is redundant with the other variables. This fact is interesting since it shows that a judgement about the logical relation between two sentences partly sums up the contribution of factors of non logical nature. All these results confirm that the notion of causality includes relations of various natures, which all participate to the notion of causal force.

Discussion

Our results show that for the types of texts we selected, causal force has the same topological properties as probability with regard to composition in causal chains. This study must be completed with an experiment measuring reading times, which would allow for the reproduction of Keenan, Baillet and Brown's (1984) results, for texts instead of just couples of sentences. Indeed, if a reading time experiment reproduces the same patterns as those that we found in this study, this would confirm causal force as a reliable predictor of reading times, as well as it would enrich the current multiple-indexing theory of text comprehension, by a description of the causal dimension richer than just the continuity/discontinuity opposition: Causality, as described by causal force, has various degrees, and is composed of variables which are not only logical.

Comparing the Contribution of the Visuo-Spatial Sketchpad and Phonological Loop in the Written Recall of a Story

T. Olive, V. Li Calzi & A. Piolat

University of Provence (France)

Thierry.Olive@newsup.univ-mrs.fr

Text and story recall has been widely used to study encoding, storage and retrieval of information in long term memory. But, in the case of a written recall, performance is not only determined by the mechanisms of long-term memory. Indeed, in a written recall of a story, subjects activate retrieval processes but also effortful writing processes (see Piolat and al. (1986). Our goal in conducting this research was to investigate the effect of the working memory slave systems (Gathercole & Baddeley, 1996) involved in the writing processes when recalling by writing a story that was presented either with verbal or with figurative material. With a verbal presentation of the story, resources allocated to the writing process of Translating are minimized because linguistic and semantic information of the story are provided to the participant that can focuses its attention on retrieval of information. By contrast, during the recall from a figurative presentation, Translating is activated and writers must share their working memory resources between this writing process and retrieval of information. To test how the slave systems are involved in these two conditions of recall, we used two secondary tasks: a phonological secondary task designed to load the phonological loop (exp. 1) and a visual secondary task designed to load the visuo-spatial sketchpad (see note 1).

Experiment 1

In this experiment we expected that performance on written recall from a figurative presentation of the story should be lower than with a verbal presentation of the story because the phonological secondary task should disrupts operations of the writing process of Translating.

Participants, procedure and material. Participants were split into two experimental groups (n = 12) according to the type of presentation of the story. The semantic correspondence between the verbal and figurative versions of the story was established by Denhière and Langevin (1981) on the basis of the predicative analysis (Kintsch & van Dijk, 1978). Participants recalled by writing the story while being simultaneously submitted to a phonological secondary task designed as following.

First, an alert signal informed the participant that a target item (a pseudo-word) was displayed on the screen of the monitor (for 4 sec.) and had to be memorized. About 20 sec. later, the target item was presented with two other pseudo-words (constituted with the same letters) and participants had to select the target item. About 10 s. later, a new target was presented to the participant.

Results

First, compared to the control groups, performance in dual task condition (recall and secondary task) was lower. With a verbal presentation of the story, participants recalled less sentences ($F(1, 19) = 5.722$, $p < .05$) and wrote less words per minute ($F(1, 19) = 6.006$, $p < .05$). Further, they made fewer correct answers ($F(1, 30) = 53.627$, $p < .0001$) and took more time to select the target items ($F(1, 30) = 10.322$, $p < .001$). Participants who received the figurative presentation of the story recalled significantly less sentences ($F(1, 19) = 4.768$, $p < .05$) and wrote less words per minute ($F(1, 19) = 12.937$, $p < .01$). Participants in this condition also made fewer correct answers ($F(1, 30) = 22.689$, $p < .0001$) and took more time to select the target items ($F(1, 30) = 39.042$, $p < .0001$). Second, compared to a verbal presentation of the story, participants recalled less sentences ($F(1, 22) = 6.144$; $p < .05$) with a figurative presentation of the story.

Discussion

Compared to a condition where participants only have to recall the story, a lower performance (both on the recall and on the secondary task) is observed when participants recall and perform simultaneously a phonological task. This result support the idea that with a written recall, the phonological loop of working memory is involved because the process of Translating is activated (Kellogg, 1996; Levy, 1999). As expected, the larger interference of the phonological task observed when the written recall was made from a figurative presentation of the story indicates that this later mode of presentation is more resources demanding than a verbal presentation.

Experiment 2

The objective of this second experiment was to test whether the visuo-spatial sketchpad (VSS) was also involved in the recall of the story. Consequently, we used a visual secondary task.

Participants, procedure and material. Participants were split into two experimental groups (n = 12) according to the presentation of the story. They recalled by writing the story while being simultaneously submitted to a visual secondary task. The procedure was the same as in the first experiment except that in the visual secondary task the target item was the font of a non-words and that participants had to select the target font between three fonts of three new non-words.

Results

First, compared to the control groups, in dual task condition and with a verbal presentation of the story, participants recalled less sentences ($F(1, 19) = 4.189$, $p = .0548$) but did not differ in the

number of less words per minute they produced. They made fewer correct answers (F(1, 30) = 23.313, p < .0001) and took more time to select the target items (F(1, 30) = 7.059, p < .01). Participants who received the figurative presentation of the story also recalled significantly less sentences (F(1, 19) = 6.75, p < .05) and made fewer correct answers (F(1, 30) = 10.153, p < .001) but they did not took different time to select the target items. Second, compared to a verbal presentation of the story with a phonological secondary task, participants recalled less sentences (F(1, 22) = 17.124, p < .001) but they took less time to select the target item (F(1, 44) = 4.343, p < .05) with a figurative presentation of the story.

Discussion

As in the first experiment, compared to a situation where participants have only to recall the story, a lower performance is observed when they have to recall the story and simultaneously perform a visual secondary task. This result indicates that the VSS is also involved during the written recall of a story. However, as among all the recall variables only the number of sentences recalled was affected by the secondary task, we argue that the VSS only affected the retrieval of information and not the process of Translating. Again, participants that were submitted to the visuo-spatial secondary task with a figurative presentation recalled fewer sentences than with a verbal presentation of the story. Surprisingly, they selected more rapidly the target font.

Conclusion

Recalling a story by writing imply the activation of writing processes constrained by working memory resources. Consequently, when recalling a story that is presented verbally the phonological loop is slightly involved because lexical and syntactical information about the story are provided to the participant. However, with a figurative presentation of the story, participants must transform into language the images that constitute the story. This operation involves both the phonological loop and the visuo-spatial sketchpad.

Note 1. A preliminary study indicated that the type of presentation of the story (verbal or figurative) did not affect significantly the recall of the story in terms of number of sentences recalled, duration of the recall, number of words, and words per minute. Further, this study indicated that the performance on the secondary tasks (phonological and visual) did not differ in terms of correct answer but differs in terms of time to choose the target item. Consequently, we conducted the analysis of performance on the secondary tasks with the percentage of decrease of correct answer and of increase of choice time.

Diagnosis of Cognitive Functioning During Text Reading and Comprehension by Adolescents

H. Thomas * & G. Denhière**

*IFAC/CFA(France)

**University of Provence (France)

thomas@ifac.cci-brest.fr

Before deciding to propose learning methods and remediation technics for disabled comprehenders, it's necessary to precisely diagnose the underlying cognitive structures and processes involved in the observed performances deficit. Diagnosis of cognitive functioning during reading-for-understanding differs radically from simple test administration. Results obtained on different tasks used in diagnosis give us an indication of the origin of a person's difficulties, and allow us to ascribe his final performances to the cognitive structures and processes at play in these tasks. Above all, the findings indicate the state of semantic information at each of the three main phases of processing : Initial text processing resulting in the step-by-step construction of the text's meaning; Memory storage of the meaning constructed ; and Retrieval and production of stored information in compliance with the constraints of the task (Denhière, Marouby-Terriou & Tapiero, 1994). For a given overall comprehension level, our aim is first to assign profiles to persons by identifying the sources of the observed difficulties and successes, and second, suggest appropriate training in accordance with the cause of the main difficulties observed during one of the three phases of semantic information processing : Input, storage, and retrieval. With this aim we used DIAGNOS-LECTURE, an alternative reading test to the standard ones.

Method

Material. Sixteen stories containing the same number of words, propositions and sentences but differing in the number of episodes were used. Judgments of the relative importance of information in the 16 stories obtained from adult subjects were used for comparison with children and adolescents. The number of sentences to order in the scrambled stories varied from four to eight. For half of the stories, overall cohesion was organized in terms of causality in the physical world, and for the other half, in terms of causal intentionality (Thomas, 1999).

Subjects. Fifty-six subjects (16 to 20 years old) participated in this experiment. They were apprentices at the IFAC Institute of Brest.

Tasks. Subjects had to perform the following tasks:

Input phase: Vocabulary test, judgment of relative information importance in n stories, ordering scrambled stories, and summarizing a complex sentence in a single word.

Output phase: Free recall of n stories, cued recall, recognition, and summarizing.

Storage phase: Immediate and delayed (15 days) free and cued, recall and immediate and delayed (15 days) recognition.

Results

Based on previous results (Denhière, Thomas, Legros & Caillies, 1997) this cohort of 56 subjects could be categorized into 5 profiles.

Profile 1: Input difficulties. The reading times of these subjects did not depend on the relative importance of the information. The importance judgments of some of these subjects were consistent with adult judgments.

Profile 2: Storage problems. For these subjects, the difference between immediate and delayed recall performance was dramatic. A large number of subjects were unable to recall anything about the story read fifteen days earlier.

Profile 3: Loss of surface information. While immediate and delayed recall performance were acceptable, recognition performances was poor on old and new information.

Profile 4: Semantic information retrieval problems. These subjects had poor immediate free recall performance, and substantial difficulty in retrieval of semantic information when cued sentences were presented and when causal questions were asked after recall (Graesser & Clark, 1985).

Profile 5: Low cognitive efficiency. Taking the input and output levels together, we computed a cognitive efficiency ratio equal to the number of recalled propositions (identical and similar to the base text) divided by the reading time per proposition.

Conclusion

The main results of this study further validate the five disabled-reader profiles. These five profiles have been found in the results of over 150 children and adolescents. In order to promote better comprehension, three solutions are possible: Modify the characteristics of texts (Britton & Gulgoz, 1991), try to change the behavior of the reader or both. What is really new with DIAGNOS then, is that depending on a person's profile, specific learning methods and remediation techniques can be developed and tested. Through ongoing interaction between cognitive psychology and learning and remediation at school, we can test our theoretical views as well as help teachers and speech therapists (Kintsch & Kintsch, 1991). In the future, the use of DIAGNOS could be considerably facilitated by applying the Latent Semantic Analysis method (LSA) proposed by Landauer and Dumais (1997). The first step towards progress in this direction would be to compare previously analyzed recall and summary protocols, with the cosine data output by LSA.

References

Britton, B. K., & Gulgoz, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. Journal of Educational Psychology, 83, 329-345.

Denhiere, G. & Baudet, S. (1990). Le diagnostic du fonctionnement cognitif dans la compréhension de textes. In Glossa, les cahiers de l'Unadrio, 20, 10-17.

Denhiere, G., Marouby-Terriou, G. & Tapiero, I. (1994). Le savoir-lire et le non-savoir-lire : diagnostic, simulation et remédiation. In J. Y. Boyer (Ed.), Evaluer le savoir-lire (pp. 199-243). Montréal, Les Editions Logiques.

Denhiere, G., Thomas, H., Legros, D. & Caillies, S. (1997). L'identification des profils d'apprentis en difficulté à l'aide de la batterie Diagnos Lecture. In V. Espérandieu (Ed.), Les troubles du langage (pp. 65-77). Paris, Ministère de l'emploi et de la solidarité.

Graesser, A. C., & Clark, L. F. (1985). Structures and procedures of implicit knowledge. Norwood, NJ : Ablex.

Kintsch, E. & Kintsch, W. (1991). La compréhension de textes et l'apprentissage à partir de textes : la théorie peut-elle aider l'enseignement ? In Les entretiens Nathan, La lecture, Actes 1 (pp.13-28). Paris: Nathan.

Landauer, T. K. & Dumais, S. (1997). A solution to Plato's problem : the Latent Semantic Analysis. Theory of the acquisition, induction and representation of knowledge. Psychological Review, 104, 211-240.

Thomas, H. (1999). Diagnostic du fonctionnement cognitif, évaluation dynamique et remédiation. Pour une approche interdisciplinaire. Thèse de doctorat nouveau régime. Université de Rennes 2.

Representation and Processing of Contrastive Information in discourse Comprehension

J. M. Lee, S. Choi, J. G. Lee, & K. H. Cho

Sung Kyun Kwan University (South Korea)

jmlee@yurim.skku.ac.fr

Eight experiments were conducted to examine in what ways CONTRASTIVE information in a discourse has positive effects on the formation of a coherent representation of a discourse and the subsequent memory of it. A three-process model was proposed. The processing of contrastive information in a discourse could involve; 1) the process of finding the nature or levels of mismatch between the current input information and the mental model or the situation model already formed based on the previous information given in the discourse, 2) the process of resolving the mismatch through activating new additional relevant knowledge structures and generating a series of inferences, and 3) the process of integrating all the relevant information into a coherent and higher abstraction level representation of the discourse (Lee, 1979). These additional processing stages could entail longer processing time at the time of encoding. In the final representation, the contrastiveness itself could be stored as a higher abstraction level proposition that will serve as an effective higher-level retrieval cue and shorten the retrieval time for the contrastive information. It could be further assumed that the ease of processing contrastive information will be dependent upon the contrast levels in the representational structure of the discourse. If the mismatch is at the global structural level (the macroproposition level; Kintsch & van Dijk, 1978), it will necessitate an overall and higher level reorganization of the previously formed representation. This will entail more extensive and deeper processing (longer encoding time), and in turn lead to a more elaborate and higher level representation which will be easier to access at retrieval. On the other hand, if the contrast was brought about by a mismatch at the local structural level (the micro-propositional level in Kintsch's term), this information will require minimal processing, thus be encoded faster, and yet accessed less fast at retrieval. To test these assumptions eight experiments were conducted.

Experiments

Experiment 1 (reading time), 2 (sentence-cued recall), and 3 (free recall) were conducted to investigate the processing characteristics of contrast information present in short discourses. Two sets of thirty-two short discourses (four sentences long) were used. The results showed that the CONTRAST (the fourth) sentences took significantly longer to encode than the NONCONTRAST

ones, and yet were recalled in greater amount. In Experiment 4, 5, and 6, the effects of contrast levels in discourse (fifteen sentences long) structure upon reading time and primed recognition time were investigated. In Experiment 4, the sentence next to the last one in each discourse was designated as the target sentence. The target sentence was constructed as having CONTRASTive or NONCONTRASTive relation to other preceding sentences. The target sentences were of either a macrostructure sentence or a microstructure sentence in Kintsch's term. Experiment 5 was a replication of Experiment 4, with script structure added. The general results of these two experiments showed a statistically significant interaction effect between contrastiveness and contrast levels. With NONCONTRAST sentences, macro- and micro-sentences were processed in almost the same amount of time, while with CONTRAST sentences macro-sentences took longer to process than for micro-sentences. The same trend was observed for the sentences immediately following the target sentences. In the retrieval test (primed recognition test), CONTRAST sentences were recognized faster than NONCONTRAST ones, and contrast levels had effects for NONCONTRAST sentences only. Macro-sentences took longer to be recognized than micro-sentences in NONCONTRAST condition only. In CONTRAST condition, recognition of macro- and micro-sentences took almost the same amount of time. In Experiment 6, the resolution of the mismatch (conflict) brought about by the CONTRAST sentence was varied as follows. The sentence immediately following the target sentence did (RESOLUTION: RS) or did not (NONRESOLUTION: NR) carry the information that explained, elaborated, or resolved the mismatch (contrast). The results of the experiment showed that the reading time of the last sentence was not significantly different between RS and NR conditions. A contrast level effect only showed up. It seems that the sentence following the CONTRAST sentence is processed sufficiently long, to integrate it with the previously formed representation, regardless of whether it does or does not resolve the mismatch. It appears then that certain attempts to resolve the contrast or mismatch are made before the onset of the next (resolving) sentence. The positive effect of processing a RS sentence appeared at retrieval only; RS sentences were retrieved in greater amount (more than twice) and faster than NR sentences. The contrast levels effect showed up only in the retrieval test: Macro-sentences were retrieved faster than micro-sentences in the NR condition only. Experiment 7 examined, by presenting the target sentences randomly (devoid of the story or script discourse context), the possibility that the above effect was due to some uncontrolled confounding effect of materials. The results showed no reading time differences across conditions. Experiment 8 was conducted, employing different types of story structures (linear and hierarchical (story grammar) structure). The general trends of the other experiments were reconfirmed: longer encoding time but shorter recognition latency for CONTRAST information (regardless of the types of story structure). This effect was stronger with the macro-level contrast.

General Discussion

The general results of the eight experiments were interpreted as supporting our research hypotheses; that the contrast information in a discourse could be represented as a higher macroproposition, and that it makes the discourse more coherent at a higher abstraction level. It seems we need a more comprehensive concept of coherence in discourse comprehension that includes the contrast information processing as an upper-level-coherence construction device.

References

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of discourse comprehension and production. Psychological Review, 85, 363-394.

Lee, J. M. (1979). Deeper processing: Spreading elaboration and integrative elaboration. Unpublished Ph. D. Dissertation. Queen's University (Canada).

Inter-Individual Differences in the Integration Processes of Text Comprehension: Do they Influence the Use of the Leading Edge Strategy and Comprehension?

L. Demanet, M. A. Schelstraete & M. Hupet

University catholic of Louvain (Belgium)

demanet@exco.ucl.ac.be

Almost every model of language comprehension postulates integration processes accounting for how new information is incorporated with already known information (e.g. Kintsch, 1998, Gernsbacher, 1990). In their first model of text comprehension, Kintsch and van Dijk (1978, 1983) postulated that certain parts of an already-built text representation have to be maintained in a working memory buffer in order to guide the integration of new propositions in the long term-representation of the text. They further postulated that the information maintained in this buffer resulted from encoding strategies, the most significant of which being the so-called "leading edge strategy" (hereafter L.E.S.). This strategy leads to select and to store in the buffer the most important and the most recent propositions of the currently read materials. More recently, Kintsch (1998) slightly modified this conception by proposing that the selection of information would be determined only by the importance parameter. Moreover, the content of the buffer is supposed to be regularly updated because of its limited capacity. The content of the buffer must then be integrated with the previous information about the text in a long-term working memory (WM-LT) to obtain a coherent representation of the text. In his experiment, Fletcher (1981) confirmed the use of L.E.S in showing that young people recognized more accurately but also more quickly propositions retained in the buffer in comparison with other propositions. The present experiment aimed to find out to what extent inter-individual differences in integration capabilities in the long-term working memory were likely to condition the use of the leading edge strategy during text comprehension and to what extent they might predict the comprehension abilities.

Method

Twenty-eight young-old participants from 60 to 69 years and twenty-nine old-old participants over 75 years volunteered to participate in a Reading-Recognition-Comprehension task (see below). The experiment aimed to test the use the readers made of the leading edge strategy as well as their comprehension performances. Participants were submitted to a self-paced reading task of twelve short expository texts. The eight target texts (the other four served as fillers) were comparable in length (words and sentences) and in number of propositions. A propositional analysis (Miller & Kintsch, 1980; Bovair & Kieras, 1985; Kintsch, 1998) was used to determine which propositions would be maintained in the buffer and which ones would not be selected by the L.E.S. The participants were warned that their reading would be sometimes interrupted and that they would be tested on some information about the text they had just read. Whenever such an interruption occurred, they were actually presented with five short sentences and, for each of them, they were required to decide as quickly as possible whether it conveyed or not information that had been presented in the text they had just read. The five sentences were the following ones: A proposition that was assumed to be retained in the buffer from the first processing cycle, a proposition that was assumed to be retained in the buffer from the second processing cycle, a proposition that was assumed not to be retained in the buffer from the first processing cycle, a proposition that was

assumed not to be retained in the buffer from the second processing cycle, and the superordinate proposition.

Reading times by sentence were recorded as well as the response latencies and error rates in the recognition task. Both measures were used to test the content of the buffer when the recognition task was performed: If the reader actually applied the leading strategy, the propositions retained in the buffer were expected to be recognized quicker and more accurately than the other propositions, as observed by Fletcher (1981). After the recognition task, the reading continued at the point where it had been interrupted. Finally, the reading of each text was immediately followed by a comprehension task. Four questions were asked orally and the participants had to answer them orally. Their responses were recorded. Two of the four questions concerned the information that the participants had been presented within the recognition task. For example (translated from French):

Reading (3 sentences): The ozone layer thickness is now significantly decreasing. It is necessary around the planet because it filters the sunbeam. The ozone layer represents a shield for the skin and prevents also the sun to damage the vegetation.

Recognition task: 1/ The ozone layer thickness disappeared: NO (Key press) Or: The ozone layer thickness is decreasing: YES (Key press); 2/ The ozone layer is decreasing progressively: NO (Key press) Or: The ozone layer is decreasing significantly: YES (Key press) 3/ 4/ 5/ other test items.

Reading (3 sentences): On earth, the ozone irritates the eyes and involves respiratory troubles. Its accumulation is supported by fumes from the cars that are accumulated on the roads. At this moment, the ozone rate in the air is frightening particularly the persons responsible for the big towns.

Comprehension questions: (Q1) What's the ozone layer problem like? ; (Q2) Why is the ozone layer necessary ? ; (Q3) How could ozone be dangerous? ; (Q4) Why the ozone rate increasing in towns?

In addition to the Reading-Recognition-Comprehension task, participants were submitted to a new verbal integration task designed by Hupet, Schelstraete, Demanet and Pourtois (2000, revised manuscript). This task was inspired from the Brandsford and Franks (1971) paradigm of linguistic abstraction. Participants had to build different short stories from simple sentences they read one at a time. The stories had to be built in parallel and their number increased gradually from 2 to 4, so as the number of sentences that had to be used to build the stories. A recall score and an integration score were both computed: The recall score corresponded to the total number of simple sentences which were correctly recalled; the integration score took into account the number of correct links that can be made between the various pieces of information for a given story. For example, two stories with four sentences:

Code: A1 = first sentence about first story; B1 = first sentence about second story; A2 = second sentence about first story; B2 = second sentence about second story; A3 = third sentence about first story; B3 = third sentence about second story; A4 = fourth sentence about first story; B4 = fourth sentence about second story

(A1) La bougie s'est consumée; (B1) Sophie regarde l'aquarelle; (A2) La bougie a sali la nappe; (B2) L'aquarelle représente un manoir; (A3) La nappe était en dentelle; (B3) Le manoir a un jardin; (A4) La nappe venait d'être lavée; (B4) Le jardin est bordé de lilas

Results and Discussion

Previous experiments in our laboratory with the verbal integration task showed that some inter-individual differences could be observed among participants with regard to their integration scores, leading to the conclusion that integration abilities could differ among readers.

At present, data were just collected and are analysed in the following way: Results at the Reading-Recognition-Comprehension task are linked with the individual integration scores in order to see if better integrators will also be both better users of the leading edge strategy and better comprehenders. On a theoretical ground, results will be discussed in the context of the role of working memory in individual differences in text comprehension. Moreover, we are currently comparing our own selection of most important propositions with the simulation of the Construction-Integration model from Kintsch (1988).

Representation of Simple Arguments During Reading

A. Britt*, J. F. Rouet** & C. A. Perfetti***

*Northern Illinois University (U.S.A.)

**University of Poitiers (France)

***University of Pittsburgh (U.S.A.)

britt@niu.edu

The ability to produce and comprehend arguments is an essential element of literacy. Argumentation is a major component of social interchange, and having students engage in argumentative discourse is one of the best vehicles for producing learning and conceptual change (Piaget, 1932; Murray, 1972; Resnick, 1991; Dunn, 1989, 1992; Stein & Miller, 1995). Engaging in argumentation, whether by comprehending another's argument or by producing one's own, presupposes the ability to internally represent the linguistic information and discourse structure of arguments. While there is a considerable body of research on how people evaluate written arguments, comparatively little is known about how people comprehend and represent arguments while they are reading. Conversely, much of what we know about the comprehension of texts is based on studies of narratives (Mandler & Johnson, 1977; Omanson, 1982; Stein & Glenn, 1979; Thorndyke, 1977; Trabasso & van den Broek, 1985), with less known about how readers represent persuasive prose. We suggest that just as people acquire representational structures, or schemata, for narrative texts, people must also learn to recognize standard argument structures in order to comprehend argumentative texts. In earlier work by Marron, Britt and Perfetti (1995), we found evidence that about one quarter of American college students were skilled at both detecting and representing argument information, about one half of the students were able to represent argument information but did not do so spontaneously, and the remaining quarter of the students did not appear to be able to either detect or represent written arguments. Students who did not represent the information as an argument tended to organize their recalls of the texts in narrative form. Although these results are interesting, they do not tell us much about the structure of argument representations constructed during reading. The following experiments were designed to test whether an argument schema guides the on-line reading of simple arguments. We provided students with arguments in a standard order (i.e., claim-reason arguments) or a non-standard order (i.e., reason-claim arguments). We hypothesized that if people use a hierarchical schema to understand and organize argument information during reading, then the same argument will be read faster if information is presented in an order consistent with filling the hierachical structure than in an inconsistent order. We further hypothesized that this order is one in which a claim would be preferred before a reason. If the reason comes first then there is no open organizing frame

available and the reader must hold it temporarily until the claim is encountered. This additional processing should be revealed as longer reading times.

Method

We had undergraduates read simple claim-reason sentences such as "Recycling should be federally mandated because recycling saves the environment." Each argument was presented in a standard claim-first order (e.g.,"Recycling should be federally mandated because recycling saves the environment.") and in a non-standard reason-first order (i.e.,"Because recycling saves the environment, recycling should be federally mandated."). Participants were instructed to read the argument at a normal rate of speed using a non-cumulative phrase-at-a-time moving window. After reading each argument, they were given a related reason, such as "We are running out of available landfill space", and asked to judge whether this reason supported the same side of the argument or an opposing side of the argument. In experiment 2, participants were also asked to complete a cued recall task in which they were given a target word presented once in the claim and once in the reason (e.g., "recycling" cued the argument above). To ensure that any difference between argument segment orders is not a result of the reason being syntactically a subordinate clause, we manipulated the explicitness of the connection between the argument elements (i.e., because) in experiment 2. Half of the participants received the arguments with the marker present (essentially a replication of experiment 1) and half received the arguments without the marker (e.g., Recycling should be federally mandated. Recycling saves the environment.). In this marker absent condition, the reader has to semantically determine the relationship between the two syntactically equivalent argument elements.

Results and Discussion

Reading times. In both experiments, we found that arguments are read faster when presented in a standard claim-first order than when the same arguments are presented in a non-standard reason-first order. Furthermore, the presence of an explicit marker did not account for these results. Regardless of marker presence or absence, the arguments were read faster when the claim comes first followed by the reason. This supports our hypothesis that readers use a hierarchical schema to organize the information during the reading of an argument. A more detailed analysis of which component is most influenced by presentation order further supports the importance of the claim during reading. We found that both claims and reasons were read slower in the first position than the second position but that the difference between first and second position is greater for reasons then for claims.

Recall results. In experiment 2, the participants final task was to recall as many of the arguments as possible given a listing of target words from the arguments. A cursory analysis suggests that students recalled more claims than reasons. We are currently doing an analysis of the order of recall of the argument components to see if there is a claim-first preference regardless of presentation order.

Summary. The results of both experiments suggests that at least some readers use an argument schema during reading and later recall of arguments. We found that reading times were longer when argument information was presented in a non-standard order (i.e., reason-claim) and that claims were recalled better than reasons.

Semantic Macrostructure Effect in Accessing Relevant Prior Information during Text Reading

S. Montoya*, G. Denhière* & T. Baccino**

*University of Provence (France)

**University of Nice Sophia-Antipolis (France)

montoya@newsup.univ-mrs.fr

This study takes place in the Memory Based Text Processing (MBP) framework (McKoon, Gerrig, & Greene, 1996). This view assumes a fast-acting parallel resonance process as the critical mechanism for retrieving prior information. Myers and O'Brien (1998) proposed a model of the resonance process. According to this view, concepts and propositions derived from the text representation resonate in response to related concepts in the currently processed sentence. According to the resonance model, the final activation value of a concept depends on the structure of the semantic representation. As a result, if a given concept [a] has a higher activation value than that of the concept [b], the latest decreases in activation. Two eyetracking experiments were designed to test the hypothesis that the semantic macrostructure of text affects the accessibility of specific information. The subjects' task was to read texts while their eye movements were monitored and then to answer to a comprehension question. We manipulated (i) the semantic macrostructure of text constructing two versions (Continuation and Interruption of the action), (ii) the presence or not of a potential candidate for the pronoun, and (iii) the Location of potential candidate or no potential candidate between the two experiments. In the first experiment, the potential candidate or no potential candidate was always mentioned at the middle of sentence while in the second experiment, this information was always mentioned at the end of sentence (i.e., the last word of the sentence). Material was exactly the same for both experiments except the Location of the specific information (middle vs. end of sentence).

Method

The critical stimuli were 32 experimental narratives. Each text consisted of five sentences and comprised three parts. The correct antecedent of the pronoun was always explicitly mentioned in the first sentence of the text (Referent sentence). This sentence described an action that protagonist was achieving when an event occurred preventing the achievement of the action. The semantic macrostructure was manipulated by constructing two versions of the Development sentences. In the Continuation version, the second sentence returned the main action in the discourse focus and the concepts and propositions derived from the two next sentences were related to the original goal. In the Interruption version, the second sentence described the breaking of the main action and the two next sentences evoked actions related to the distractor event. The potential antecedent was always mentioned in the second sentence of text. This antecedent was a potential candidate for the pronoun on the basis of gender and number agreement. In the absence of a potential antecedent, it was replaced with a common noun. The Anaphoric sentence contained an unheralded pronoun which referred to a concept evoked in the common first sentence of texts (see McKoon, & al., 1996). The remaining of that sentence contained disambiguating information that was supposed to point to the relevant antecedent of the pronoun. We supposed that the correct referent would have a relatively high force strength of encoding in the Continuation version because this information is activated throughout the text. In contrast, in the Interruption version, the absence of information related to the referent leads to a progressive decrease in activation of this information during text reading. On this basis, we

supposed that the presence of a potential antecedent would not affect the processing time of the pronoun in the continuation version. In contrast, there would be a competition between the correct antecedent and the potential antecedent resulting in a longer processing time of the Anaphoric sentence. Moreover, we investigated how the text's surface form would affect the structure of the semantic representation of text.

Results

We investigated the time course of the anaphoric resolution processing by dividing the sentence into five zones. We realized combined analyses of two experiments. Two main results derived from the processing time of the Beginning zone. First, the presence or not of a potential antecedent interacted with the Development structure. In the Continuation version, fixation duration was longer when the potential antecedent was present although the difference was reversed in the interruption version. Second, when the potential antecedent was present, fixation duration was longer in the end condition than in the middle condition though there was no difference between the two versions of text when the antecedent was absent. This interaction was also significant during the processing time of the Before pronoun zone. Results indicated that fixation duration was longer in the middle condition than in the end condition when the potential antecedent was present although in the absence of potential antecedent, this difference was reversed. The main result from the processing time of the Pronoun zone indicated a longer processing time in the interruption version than in the continuation version.

Two main results derived from the processing time of the After pronoun zone. First, in the interruption version, fixation duration was longer when the potential antecedent was absent than when it was present although there was no difference between the two conditions in the continuation version. Second, Location factor interacted with Development structure. Results indicated that in the middle condition, fixation duration was longer in the interruption version than in the continuation version although the difference was reversed in the end condition. During the processing of the end zone, Location interacted with Antecedent factor. Results indicated that fixation duration was longer in the end than in the middle condition when the potential antecedent was present. When the potential antecedent was absent, the difference between Locations was greater.

Discussion

In line with our hypotheses, the success and speed of resonance process are highly related to the structure of information in the text's semantic representation that, in turn, affects the encoding strength of this information. The whole of these results suggests that the presence of a potential antecedent does not affect the processing time of the Anaphoric sentence in the continuation version. In the interruption version, results suggest that disambiguating information are not sufficient to point to the relevant referent. These results are consistent with the study of O' Brien, Raney, Albrecht and Rayner (1997). These authors investigated the conditions under which the resonance process failed in retrieving the antecedent of an explicit anaphor whereas the overlapping degree between the anaphor and its antecedent was high. Our results lead us to consider that the force of encoding strength of distant information plays a great part in the resonance process. Our results also take account for the effect of the characteristics of surface form in the resonance process.

Which Role the Position of Lexical Explanations Plays in Text Comprehension?

E. Jamet & O. Le Bohec

University of Haute Bretagne (France)

Eric.jamet@uhb.fr

For few years, working memory was put forward as a major factor determining text comprehension (Daneman & Carpenter, 1980) and effectiveness of instructional documents (e.g., Sweller & Chandler, 1994). It seems that there is an increasing consensus among cognitive scientists concerning the deciding role of the manner in which the task is structured (e.g., Reder & Anderson, 1980; Sweller & Chandler, 1991, 1994; Mousavi, Low & Sweller, 1995; Mayer, Bove, Bryman, Mars & Tapangco, 1996; McNamara, Kintsch, Songer & Kintsch, 1996). In accordance with Sweller (1993), we assume that many instructional procedures are not effective by not taking the characteristics of human processing system into account and, consequently, by not structuring the material in order to eliminate or reduce extraneous cognitive load (induced by nonessential mental activities). The present research comes within the framework of this research on instructional formats. More specifically, the purpose of this study, largely inspired by Yeung, Jin and Sweller (1998) is to assess the role the position of explanations of words plays in schoolbooks. The location of the explanations is supposed to be an important factor that can largely contribute to the extent to which the adjacent passages of the text are processed and learned. This hypothesis is mostly based on the works of Sweller and his colleagues about two important effects, the effect of redundancy (Sweller & Chandler, 1994; McNamara et al., 1996; Yeung, Jin & Sweller, 1998) and the split-attention effect (Ward & Sweller, 1990; Sweller & Chandler, 1991, 1992). The first effect occurs when the same information is presented twice into two different formats or when one source of information is already known by the reader and become unavailing; the second effect comes into sight when two different sources of information that need to be treated simultaneously in order to be understood are a long distant apart. In this work, we choose to check if the split-attention effect could be found during reading when lexical explanations are provided far from the text. In order to verify this assumption, we vary the distance between target words and the corresponding explanations by using three formats of presentation: The first two are the ones originally used by Yeung et al., (1998), in other words a glossary at the end of the text (a large physic separation between the targets words and the corresponding explanations) and a format with embedded explanations (in a small font under the target words). The third format is a new one increasingly used by French publishers in schoolbooks. It concerns the margin of the text. Nevertheless, the consequences of the way the information is structured can't be analyzed without taking the role of another primordial factor into account. Indeed, since the referent paper written by De Groot (1965) with regard to the cognitive factors involved in chess play, the factor that many researchers try to take into consideration in their works is the level of participants' expertise (Ericsson & Kintsch, 1995; Kalyuga, Chandler & Sweller, 1998). For example, the level of learner expertise is now largely accepted to be a primordial factor determining the intelligibility of a text. The same set of information can be both hardly or not understandable for a less knowledgeable reader and easily intelligible for a more experienced user. Many empirical evidence proved that "experts" solve problems faster and with fewer errors than "novices". From the angle of the cognitive load theory, the essential determinant of expert performances appears to be the schema defined as "a mental construct permitting problem solvers to categorize problems according to solution modes" (Sweller, Chandler, Tierney, & Cooper, 1990, p. 176).

Method

In our study, we try to define and control precisely this potential source of heterogeneity between participants. First, contrary to what was done in the reference study, we use a single cluster of students. The referent researchers have indeed compared the results of children (10 years old) with the results of adults (university students). In fact, they were conscious of this problem and tried to solve it in the two last experiences by using the same age for all the participants. They

used students of a secondary school (eighth grade, 13-14 years old) in Hong Kong. There, students were shared out to different secondary schools based partly on their achievement in a standardized assessment in the primary school for the purpose of secondary school placement (Yeung, Jin & Sweller, 1998). The authors used the standardized scores to assign the participants to the "novice" group and the "expert" group of the two last studies. In order to ensure a sample of "expert" students (with high ability in English language), only those who passed the English examinations in both seventh and eighth grades were included in the last experimentation. Here, to be as precise as possible, two types of expertise were taken into consideration: The vocabulary knowledge and the acquaintance of the field the text talks about. Indeed, according to Diakidoy (1998), we consider that possessing a good grounding in a specific area means, at least, that a substantial number of the main important words of the field are well known. Therefore, we assume that the knowledge base of the expert can be defined as the knowledge of most of the relevant concepts and the general knowledge of the target field as well. Expertise, as indicated before, is supposed to be able to coerce initial hypotheses about the unknown words by leading the reader to observe the position and function of the concept underlying the word in the network (Diakidoy, 1998). Furthermore, some experimental evidences (Tardieu, Ehrlich & Gyselinck, 1992) have shown that the same preliminary knowledge can support inferencing when the necessary links of information are missing. The four groups of fifty persons that we needed for the research were controlled according to the participants' expertise of the concerned field. By using a control condition (fifty persons from the same school level as the experimental groups), we assessed the well known and the unknown words according to this cluster of students. Owing to this result, the two kinds of words were taken into account in such a way that the level of expertise was as controlled as possible. We assume that a large spacing between the words and their explanations would provoke a split-attention effect and would contribute to a less efficient treatment of the adjacent passages. We make the assumption that the closest of the target words the explanations are, the best it is for the comprehension of the adjacent passages (Yeung, Jin & Sweller, 1998). The effects must be different with explanations of well-known words. The treatment of useless information should impose an extraneous cognitive load bringing up a less efficient processing of the adjacent passages. In our procedure, each group of fifty persons is confronted with only one format of presentation of the words' explanations.

Results and Discussion

The results are consistent with our hypothesis. When the unknown words are close to the target explanations (i.e., under the target word or in the margin near the corresponding passage), the adjacent passages are significantly better processed than in the control condition (Dunnett's test). On the contrary, the separated format (glossary) does not make the processing easier than when explanations are not shown. This result can be interpreted beyond the scope of the split-attention effect. Concerning the well-known words, we also obtain consistent results; here, with an interpretation in terms of redundancy: When the explanations are close to the words (under these words), a less efficient treatment of the adjacent passages is obtained in comparison with the other conditions (the control and the glossary condition). The intermediate formats (margin), from this point of view, provoke intermediate results: It does not significantly differ from the separated and integrated formats. Another dependent variable used here is the comprehension of the text passages that are not directly concerned by the lexical explanations. A linear contrast indicates that the intermediate format is the only one that can let readers to allocate more cognitive resources for the processing of the "neutral passages". To sum up, if we consider the text in its totality, the intermediate format of presentation (i.e., the words located in the margin) is the only one that is able to facilitate significantly the comprehension of the text in comparison with a text without any explanations (i.e., the control condition). This format seems to bypass the split-attention effect, to decrease the effect of redundancy in such a way that more cognitive resources can be used for the comprehension of the " neutral passages". Broadly speaking, the explanations in the margin – accessible but not in the way - seem to be the most efficient for a good comprehension and learning of a text.

Textuality and the Rhyming Principle in Narrative Superstructure

H. Masuda

University of Hawai'i at Hilo (U.S.A.)

masuda@hawaii.edu

In the field of theoretical linguistics, " grammaticality " refers to the well-formedness of the structure of sentence which is constructed as one of the underlying representational systems called syntax. This linguistic module of syntactic representations consists of a set of universal principles, being built in the mind and brain as the internalized capacity of natural languages. The " Checking " theory in the Minimalist Program (Chomsky 1995), for instance, explains how the grammaticality of a sentence is determined by the operation of " merger ", leaving behind only interpretable grammatical features. The following (1) and (2) provide a few examples (Those which are underlined are uninterpretable features).

(1) Eddie-[3MS] has-[Pres] departed. (2) *Us-[1PObj] has-[Pres, SNom, +n] depart-[inf].

While syntactic representations are processed by such analytical and locally-constrained principles, this paper argues that discourse representations in the Internalized Language are processed by more synthetic and globally organized principles. Hence, the term " textuality " is applied to refer to the well-formedness and cohesiveness of the structure which is larger than sentence. The present study focuses on a type of monologic discourse which is named "narrative superstructure" by van Dijk (1988, in preparation) and posits that there is a set of internalized rules and principles which underlie such suprasentential representations.

It has been long believed that the repetition and reiteration construct a fundamental textuality, functioning as the cohesion and coherence devices in narrative superstructure (Grimes 1970; Tannen 1989; Hymes 1990; Masuda 1995). The present paper follows the line suggested in the earlier studies by analyzing narrative discourses in creole languages. These languages are considered to be the archetypal patterns of human languages (Bickerton, 1995) that will shed light on the study of the nature of human behavior. Three varieties of creole languages, Guyanese Creole, Indo-Portuguese Creole, and Hawaiian Creole, are investigated and the analyses are carried out within the framework of the Narrative Representation Theory (Hymes 1981, 1990; Masuda 1998a, 1998b, 1999, 2000; NRT henceforth). The NRT model tells the fundamental hierarchical order of narrative superstructure is controlled by the quint-patterning principle, which projects five levels of external units: Interpretation units; coherency units; episode units; juncture units; and apex unit (i.e., IU, CU, EU, JU, AU). Each of the five external levels is governed by the N-nary-branching principle, constituting the sequential representations in three or more internal components: Expositions; complication; and denouement (e.g., JU[Ex-EU1][Cm-EU2][Dn-EU3], EU[Ex-CU1][Cm-CU2][Cm-CU3][Dn-CU4], CU[Ex-IU1][Ex-IU2][Cm-IU3][Cm-IU4][Dn-IU5], etc.).

Two aspects of repetition and reiteration are investigated in this study. One has to do with the number of repeated words in a single interpretation unit, and another is related to the number of occurrences of a word in the whole narrative or apex unit as illustrated in (3) and (4). (5) is a

sample narrative discourse elicited from Hawaiian Creole in which several reiterations (underscore) are found.

(3) X-word: 1-word [in wan long lain] "in one long line")

(4) Y-plets: Triplet [in lain], [in wan lain], [in wan long lain]

(5) AU JU   EU   CU1   IU1   No, da [ka get in wan long lain]

                          IU2   Samtaim [in wan long lain]

                                IU3   Meibii get about fiftii sikstii[kaz], yae?

                                "No, the car get in one long line, Sometime in one long line, Maybe get about fifty, sixty cars, right."

              CU2   IU4   AEn' den, its nat oni wan lain,

                        IU5   [Dei get sevro mo] aDa lainz,

                                IU6   [Dei get sam mokaz], tu, ae?

                                "And then, it's not only one line,

                                They get several more other lines,

                                They get some more cars, too, right?"

              CU3   IU7   AEnd, da tiimstaz is,

                        IU8   [Awlweiz bringin] ka, aend,

                                IU9   [Awlweiz teikin] aut, sii.

                                "And then, the teamsters is,

                                Always bringing car and,

                                Always taking out, see."

The following features are found in the analyses of repetition and reiteration. First of all, among 780 interpretation units (IUs) obtained from nine narrative discourses in the three creole languages, only 150 units do not contain any repetition of words. That is to say, as many as 630 interpretation units have some repeated words in them, accounting for 81% of the total. Secondly, the scope of X-word extends from single to nine (1-9) while the range of Y-plets goes from doublets to nonagon (2-11). Furthermore, it is also found that the distribution of Doublets tends to concentrate on the scope of an episode unit (49%) while that of multiplets focus more on a juncture unit (46%). These findings support the earlier claim that there is a set of holistic rules

that governs the suprasentential structures and the intricate distribution of repetition and reiteration manifests itself as a far more complex process than is commonly recognized.

Based on the above discoveries, the present study goes on to argue that the repetition and reiteration constitute a canonical pattern of discourse grammar and thus they are considered to establish a universal rule. This particular rule named the "Rhyming Principle" in this study could be cooperative yet autonomous from other linguistic principles in constructing narrative superstructure. The Rhyming Principle postulates that a well formed/cohesive narrative discourse must comply with the following set of conditions.

Condition A: Reiteration must play a major role in establishing the textuality/cohesion of narrative discourse. Condition B: In reiteration, a dual/plural-word appears more frequently than a single-word; a doublet is more frequent than a multiplet. Condition C: Episode/coherency units are controlled by doublets: the textuality of episode/coherency units is preserved by the application of doublets. Condition D: A juncture unit is controlled by multiplets; the textuality of juncture unit is preserved by the application of multiplets.

These findings suggest that the coherent properties for establishing "textuality" in discourse are quite different from the sentential conditions of "grammaticality" in syntax. Recent neuroscientific studies present a support for this claim. They claim that syntax and discourse are separable in terms of their cerebral computations since the center for the former is located in the left hemisphere of the brain while for the latter it focuses on the right (Hough 1990; Deacon, 1997; Obler & Gjerlow, 1999).

Building a Mental Model from Text and Pictures: The Role of the Visuo-Spatial Working Memory in the Memory of Pictures

V. Gyselinck & M. F Ehrlich

University of Paris V (France)

gyselinck@psycho.univ-paris5.fr

A variety of research suggests that pictures embedded in a text facilitate the construction of a mental model (see for a review, Gyselinck & Tardieu, 1999). Fewer experiments have questioned the locus of the construction, storage and integration of the representation built from texts and pictures. Using the concurrent task methodology, some recent data suggest that the integration of scientific texts and pictures makes use of the visuo-spatial sketchpad of working memory (Kruley, Sciama, & Glenberg, 1994; Gyselinck, Cornoldi, Ehrlich, Dubois, & de Beni, 1999; Gyselinck, Ehrlich, Cornoldi, de Beni, & Dubois, 1999). In addition, these studies suggest that the subjects' spatial span may play a role in the way subjects benefit from the presentation of illustrations. The aim of the present study is to question the role of the sketchpad in the integration of texts and pictures. Is the picture immediately interpreted and transformed to be integrated in some amodal representation? Or is it temporarily stored in the sketchpad before its integration into the representation? In the first case, one could imagine that the surface features of the picture should be rapidly lost, whereas in the second case, the surface features of the pictures should be easily retrievable. To examine these questions, subjects were presented texts with pictures or pictures only. Their memory of the surface features of the pictures was tested by means of the method used by Gernsbacher, Varner, and Faust (1990) and their comprehension of the phenomena described was also tested. In addition, the role of the spatial span was examined.

Method

57 undergraduates participated voluntarily to the experiment in exchange for a course credit. Six short texts, 9 sentences long, dealing with basic notions of physics were used. A series of 9 schematic pictures corresponded to each text, one picture for each sentence. The subjects had to read and understand three illustrated texts presented on a computer screen and they also had to visualize 3 series of pictures corresponding to the 3 other texts. Immediately after the presentation of one text (or a series of pictures), the subjects took a recognition test. They had to decide between two pictures, one was in the original version and the other one had its left/right orientation reversed. This was done successively with 3 of the pictures presented. Then the subjects took a comprehension test. Three paraphrase- and 3 inference-questions were presented successively in a random order, and for each question, the subjects had to choose between 3 responses. Finally, a new recognition test was given, with three new pictures. After the subjects had completed the whole presentations, they were submitted to the Corsi-blocks test measuring their spatial span, and also to a digit span test measuring their phonological memory.

Results and Discussion

Two analyses of variance were conducted on recognition data and comprehension data. In the first one, all subjects were considered. In the second analysis, two subgroups were formed, contrasting spatial span with the same mean scores in digit span and prior knowledge: 19 subjects formed the low spatial span subgroup (span from 2.5 to 4.5) and 18 subjects formed the high spatial span subgroup (span from 6 to 7.5). The analysis of recognition data showed that pictures were better recognized when they had been presented without text (74.2%) than when they were associated to the text (69.1%) ($F(1, 56) = 5.79$, $p < .05$). In addition, recognition was better immediately after the presentation (76.2%) than when the recognition test was delayed after the comprehension test (67.1%) ($F(1, 56) = 23.66$, $p < .0005$) and there was no interaction. In the second analysis, there was a marginally significant effect of the subgroups: high spatial span subjects had better recognition performance (75.2%) than low spatial span subjects (69.3%) ($F(1, 35) = 3.60$, $p = .06$), but no interaction was found. Analyses on comprehension data revealed that performance was higher when subjects were presented pictures with texts (74.8% of correct answers) than when only the pictures were presented (65%) ($F(1, 56) = 23.20$, $p <. 0005$), and there was no interaction with the type of questions. The second analysis showed that high spatial span subjects performed significantly better (74.1%) than low spatial span subjects (65.6%) ($F(1, 35) = 8.60$, $p < .01$), and there was no interaction.

Altogether these results show that the surface features of the pictures are easily retrievable, which suggests that the pictures are not immediately interpreted and transformed to be integrated in some amodal representation. They would rather be temporarily stored before their integration into the representation. When only the series of pictures are presented, subjects probably have to store them before they are able to build a coherent representation. When the pictures are presented with texts, the interpretation of the pictures could be more directly performed with the help of the corresponding sentence, and the need for holding pictures in memory would be less important. This could well be done in the visuo-spatial sketchpad of working memory as is suggested by the better recognition performance of the high spatial span subjects compared to low spatial span subjects. High spatial span subjects seem able to process the visuo-spatial material more easily, memorizing it slightly better and, more importantly, integrating and interpreting it better to be able to answer questions about the phenomena evoked.

References

Gernsbacher, M.A., Varner, K.R., & Faust, M.E. (1990). Investigating differences in general comprehension skill. Journal of Experimental Psychology: Learning, Memory, and Cognition, 16, 430-445.

Gyselinck, V., Cornoldi, C., Ehrlich, M.-F., Dubois, V., & De Beni, R. (1999). Visuospatial memory and phonological loop in processing texts and illustrations (Submitted).

Gyselinck, V., Ehrlich, M-F., Cornoldi, C., De Beni, R. & Dubois, V. (1999). Visuospatial working memory in learning from multimedia systems. Journal of Learning Computer Assisted Learning, in press.

Gyselinck, V. & Tardieu, H. (1999). The role of illustrations in text comprehension: what, when, for whom, and why? In S. R. Goldman and H. van Oostendorp (Eds.), The Construction of Mental Representations During Reading. N.J.: Lawrence Erlbaum Associates.

Kruley, P., Sciama, S.C., & Glenberg, A.M. (1994). On-line processing of textual illustrations in the visuospatial sketchpad: Evidence from dual-task studies. Memory and Cognition, 22, 261-272.

Story Comprehension in Patients with Damage to Prefrontal Cortex

T. Zalla*/, M. Phipps**, M. Colvin** & J. Grafman****

*Institute of Cognitives Science, Bron (France)

**National Institutes of Health, Maryland (U.S.A.)

zalla@isc.cnrs.fr

Patients with frontal lobe damage are reported to have difficulties in processing complex event information. In previous studies, we showed that patients with prefrontal lobe damage are selectively impaired in the maintenance of sequential and hierarchical order of information related to action (Sirigu et al., 1995; Sirigu et al., 1996). We also showed that the representation of temporally organized sequence of events requires different processes from those needed for the comprehension of individual words or well formed syntactic sentences (Sirigu et al., 1998). The present study was designed to test the hypothesis that patients with frontal lobe lesion would show difficulties in establishing the inferential structure of a story as a result of an impairment in generating causal connections between events. Patients with frontal lobe lesions are reported to be disturbed in certain cases of story comprehension (Luria, 1970). Clinical anecdotes often claim that patients with frontal lobe lesions are unable to give a detailed description of a picture or to tell a story during free-recall, but can do so when presented with probe questions to "clarify" the story (Luria, 1970). Luria (1963) showed that by having the patient read a passage and by making "appropriate remarks" in reading the story, the patient can use those remarks subsequently to retrieve the passage. Conversely, studies with aphasic patients showed that although these

patients were impaired at the level of sentence processing, they were capable of grasping the general meaning of statements and fables (Luria, 1970; Freedman-Stern et al. 1984).

The focus of this study is on story comprehension ability in patients with frontal lobe lesions or atrophy. We assessed the ability of a group of 15 patients with frontal lobe lesions, a group of 15 patients with frontal lobe dementia and a group of 7 amnesic patients to comprehend stories under two different conditions. The two experiments consisted of two different stories in which different working memory demands were required to the subjects in order to comprehend the stories. The subjects were first presented with the Story "The Czar and His Daughters" (Mandler & Johnson, 1977; Rumelhart, 1977; Graesser, Robertson & Anderson, 1981) and, in a separate testing session, the same groups of patients were presented with the Story "The Ant and the Dove" (Mehan, 1977; Graesser et al., 1981). The conceptual structure of each story consisted of a set of nodes or structural components. These node categories were adopted from Graesser et al., (1981). The question-answering (Q-A) procedure has been proven to be a valid method of investigating narrative organization (Graesser, 1978; Graesser, Robertson, Lovelace & Swinehart, 1980; Graesser, Robertson & Anderson, 1981). This procedure enables investigators to elicit explicit and implicit information that subjects inferentially extract from the story during comprehension. Two different Q-A procedures were applied in the two experiments. In the first experiment, general (i.e., "what type of story was this?") and specific questions (i.e., "Which characters were first introduced?") were asked after free recall of Story I (approximately 60-90 minutes later). Only in the experiment 2, were questions asked in order to test story comprehension during story presentation. In this condition, Q-A procedures allowed us to verify the capacity of the subjects to establish sequential causal connections on-line during the process of comprehension along with the ability to construct inferences regarding information that is not immediately available in the text.

Our results show that patients with frontal lobe damage and frontal lobe dementia have severe difficulties in establishing and retrieving the inferential structure of a sequence of story events in both conditions while they could recall the individual events (words and individual propositions) outside their narrative context. Conversely, amnesic patients were impaired in retrieving the single episodes. We further examined the effects of lateralized lesion location on performance in the FLL group. We divided the frontal lobe patients into three groups on the basis of the localization of the lesion: One group of 5 patients with left frontal lobe damage, 4 patients with right frontal lobe damage, and 4 patients with bilateral prefrontal lobe damage. Different patterns of deficits in story comprehension were found with respect to the localization of the lesion. Patients with damage to the right prefrontal cortex appear to be significantly less impaired than patients with bilateral and left prefrontal lesions in both conditions. Performance on story recall measures was compared with scores representing intellectual ability, degree of dementia and mood/social conduct. The selective impairment in retrieving the structural components of a story observed in patients with frontal lobe lesions would account for their inability to form the overall meaning of the story. Our results also indicate that the left prefrontal cortex would be more involved in story comprehension because of the linguistic and propositional nature of the representations.

Functional Hemispheric Asymmetry in Voicing Feature Processing in Reading

N. Bedoin & H. Chavand

University of Lyon 2 (France)

bedoin@univ-lyon2.fr

In previous research, we assessed whether the phonological code involved in early stages of print processes was organised in terms of phonetic features. Infra-phonemic resemblance between printed stimuli was manipulated in experiments conducted in French (Bedoin, 1998a).

In Experiments 1-5, the phonemic similarity between prime word and target word was kept equivalent across conditions, but the phonetic overlap varied. In the high similarity condition (HR), the difference between prime and target was restricted to place and/or manner of articulation features of the initial consonants (e.g. passe-TASSE); in the low similarity condition (LR) there was an additional difference in voicing feature (e.g. basse-TASSE). Lexical decision latency was longer in HR than in LR, whatever the SOA (100, 66, 33 ms), whatever the prime word frequency (i.e., more or less frequent than the target) and whatever the prime was a word or a pseudo-word. These data provided evidence of a fast, pre-lexical and automatic activation of phonetic properties in silent reading.

In Experiment 6 (Bedoin, 1998b), the voicing resemblance effect was tested with the backward visual masking procedure (Naish, 1980). The target was briefly presented at first (33 ms) and stimuli used as prime in Experiment 5 were used as masks (33 ms). Subjects had to recall the target. Other researches showed that the disruptive effect of the mask was reduced if there was a graphemic overlap with the target. A similar mask was supposed to reinstate partial information extracted from the first stimulus, allowing better identification than a control mask. It has been consistently found that a phonemic overlap further reduced the masking effect (Perfetti, Bell, & Delaney, 1988; Perfetti & Bell, 1991). Experiment 6 showed that a phonetic overlap also reduced the masking effect and supported the idea that phonetic features are crucial units in early phonological code. Various kinds of mechanisms are proposed to explain theses data. Whatever interpretation selected, they provided converging evidence that assembled phonology in reading is fine enough to provide very accurate phonetic distinctions.

We extended this research to assess the role of other phonetic features: The hypothesis was that the feature's class (i.e., voicing, place and manner) could lead to distinct effects, since they have been described as separate hierarchical feature tiers in some linguistic models of non-linear phonology (Clements, 1985). Further priming experiments have been conducted, manipulating place and manner similarity (Bedoin & Chavand, 1998; Chavand & Bedoin, 1999). Taken together, the data showed slight differences in place and manner features processing, providing support for psychological reality of feature categories and their geometrical organisation. Furthermore, interaction between categories appeared: Place and manner effects varied with voicing context.

Provided that phonetic feature categories seem to have a psychological status in reading we addressed the question of their neuropsychological status. In Experiment 9, we manipulated voicing similarity between printed CV target (presented during 68 ms) and printed CV mask (presented during 68 ms). Subjects were required to recall the target. Both target and mask were presented either in the left visual field (right hemisphere) or in the right visual field (left hemisphere). This visual-half-field presentation was used to examine the accuracy of target processing as a function of which hemisphere directly received the stimulus and initiated processing. Initial consonants in target and mask were always different but the vowel was always the same (U). Target and mask could be similar in voicing feature (both initiated by voiced consonants or both initiated by voiceless consonants) or different in voicing feature (one initiated by a voiced consonant and the other by a voiceless one). Targets were printed in upper case and masks in lower case. Half of the target-mask stimuli were initiated by fricative consonants and the others began with a plosive.

Results argued for the influence of phonetic manipulation on reading performance. Moreover, voicing manipulation did not affect performance the same way in each visual field. In the right visual field (devoted to test more precisely the left hemisphere involvement in processing) no effect occurred. In the left visual field (devoted to test more precisely the right hemisphere involvement in processing) the voicing manipulation significantly affected performances. Therefore, the right hemisphere revealed specifically involved in voicing feature processing during reading.

Some research has provided evidence supporting the position that speech perception mechanisms at feature level are distributed asymmetrically across the hemispheres. Dichotic studies showed that the right hemisphere might play a special role in categorical processing of voicing (Simms, Mollifies, & Brendan, 1997). Since the pattern of results obtained in reading was congruent with such a lateralised effect, our data supported the idea of common neuro-functional basis for phonetic features units involved both in speech perception and in reading. Moreover, they also evidenced the neuropsychological status of phonetic features in reading. Finally, similarity in voicing produced contrastive effects in voiced versus voiceless target-mask stimuli. Voicing similarity increased accuracy in voiced consonant recall and reduced accuracy in voiceless consonant recall. So, phonetic resemblance can help or hurt the perception of print, depending on refined aspects of phonetic structure.

"My Recaller is on Vacation"

Can memory books change discourse patterns in residents with dementia?

K. Dijkstra*, M. Bourgeois*, G. Petrie*, L. Burgio** & R. Allen-Burge**

*Florida State University (U.S.A.)

**University of Alabama (U.S.A.)

dykstra@darwin.psy.fsu.edu

We will discuss results of a discourse analysis of conversations between nursing assistants and nursing home residents with dementia before and after the implementation of an intervention program. This intervention program included a communication-training program for nursing assistants and the use of " memory books " with residents. The objectives of this study were to make a detailed categorization of discourse patterns of residents in different stages of dementia, to assess differences in discourse patterns before and after implementation of the intervention program, and to explain these discourse patterns in the light of current theories on discourse comprehension and aging. Discourse analysis research of conversation samples of persons with cognitive impairment is relevant in three respects. First, discourse produced by persons with dementia is fundamentally different from that of persons without dementia. Discourse of persons in different stages of dementia is less cohesive, coherent, and informative than that of persons without cognitive impairments. Secondly, discourse analysis of persons in different stages of dementia can contribute to a more detailed categorization of discourse patterns. Common distinctions of dementia in early-stage, middle-stage, and late-stage can be related to the discourse characteristics of the patients. Thirdly, the study of discourse patterns among dementia patients can contribute to the discussion of mechanisms explaining age-related and cognitive deficits in discourse processing.

The discourse analysis schema included discourse characteristics at a macro or discourse level and a micro or utterance level of both conversation partners (resident and nursing assistant). The macro-level included utterances, information units, words, unique words, and global coherence. The micro-level included discourse categories, such as referential and temporal cohesion, aborted phrases, empty phrases, indefinite terms and local coherence. Discourse categories of the nursing assistants included facilitators (repetitions, encouragements, and cues), questions and prompts, all

at the micro-level. An intervention study aimed at facilitating communicative interactions between nursing assistants and residents with the help of memory books was conducted in seven nursing homes in Florida, USA, to examine potential effects of this program on changes in discourse patterns among these residents. The intervention program included a three-week training program for nursing assistants that taught them to communicate more effectively with memory impaired residents and to use personalized memory books with them. These memory books were 12-page laminated booklets with information about the resident's life experiences, family, and current activities in the nursing home. They were meant to help the residents remember relevant information about their past and present life. It was hypothesized that deficits in coherent, and cohesive discourse could be minimized or countered as a result of the implementation of effective communication techniques and use of memory books.

A total of 125 residents in seven long-term care facilities in Tallahassee and their assigned nursing assistants from the morning or evening shifts participated in a 3-month intervention study. Half of the nursing assistants-residents dyads were assigned to a treatment group in which the nursing assistants participated in the training program and the residents received the memory books and half the dyads were assigned to a control group that did not receive training or memory books. Before and after the communication training, 5-minute conversations between the resident and their assigned nursing assistants in treatment and control groups were audio- and videotaped. Sixty pre- and sixty post-training conversations between nursing assistants and their residents were selected for analysis. The selection was made on the basis of the following criteria: 1) same resident-nursing aide dyad, 2) equal division of dyads in the treatment and control group (30 each), and 3) equal representation of different stages in dementia (10 each: Early, mid, late stage dementia). Early, middle, and late stage dementia was determined by the Mini-Mental State Exam score (maximum score = 30). A score of 18-25 was associated with early stage dementia, 10-17 with middle stage dementia, and 0-9 with late stage dementia. The average MMSE score in the three selected groups was 20 (treatment, SD = 2.2), 21 (control, SD = 2.2) for early stage, 13 (treatment, SD = 1.8, control, SD = 2.4) for middle stage, and 4 (treatment, SD = 2.6) and 6 (control, SD = 2.7) for late stage, respectively. Mean age of the groups was 85 years (treatment, SD = 5.4 and 85 (control, SD = 7.1). Educational attainment was 13 years (treatment, SD = 3.5) and 13 (control, SD = 3.7).

The transcribed conversations were interview-style conversations between the resident and their nursing assistant in which the nursing assistant asked the resident three questions: 1) "Tell me about your family", 2) "Tell me about your life", 3) "Tell me about your day". The nursing assistant was prompted by the research team to stick to those questions and only provide prompts, such as "tell me more" when the resident did not respond. This instruction was given in order to avoid speaker dominance by the nursing assistant. The conversations were transcribed and analyzed according to the discourse analysis schema discussed above.

The hypotheses were: 1) Distinct discourse patterns can be demonstrated among early, middle- and late-stage residents with dementia; 2) information content, coherence, and cohesion, will be higher for residents with memory books than for residents without memory books; 3) discourse of trained nursing aides will be more facilitative than that of untrained nursing aides as a result of the communication training.

In general, the results corroborated the hypotheses. First, distinct discourse patterns were established according to the stage of dementia. Among others, coherence and cohesion of discourse were lower in the middle and late stage dementia than in the early stage. Secondly, discourse patterns changed as a result of using memory books. These changes were more prevalent among the middle stage and late-stage residents with dementia in the treatment group than among early-stage residents with dementia. Middle stage residents with dementia in the treatment group had higher globally coherent discourse, and a smaller number of empty phrases than their counterparts in the control group. Late stage residents with dementia in the treatment group had higher locally and globally coherent discourse, higher referentially cohesive discourse, more unique words and a higher number of information units than their counterparts in the control

group, as well as a lower number of disruptive topic shifts and indefinite terms. Thirdly, trained nursing assistants tended to be more facilitative in their conversation than untrained nursing assistants, and provided more cues for late stage dementia residents when they encountered word-finding difficulty. The intervention program appeared to have positive effects on the discourse of middle and especially late stage residents with dementia. The use of memory books and facilitative communication techniques helped residents to remain focused on the topic of conversation and provide information to a greater extent than in control conditions. Memory books and facilitative speech may help keep information activated longer in working memory and be used in further research on programs to improve cognitive functioning among persons with dementia.

The Effect of Context and Categorizability on Verbal Learning

T. Guthke*, E.C. Ferstl** & A. Hauptmann*

*University of Leipzig (Germany)

**Max-Planck-Institut of Cognitive Neuroscience, Leipzig (Germany)

guthke@cns.mpg.de

The investigation of text comprehension in brain damaged patients is important for theoretical and clinical reasons. In this study we looked at the interaction of language and memory functions. In the clinic we see different kinds of language disturbances: Aphasic patients have especially problems on the word and sentence level, but brain injury often results in deficits of text level processes without traditional " aphasic symptoms ". We see these deficits very often in patients with traumatic brain injury or after focal lesions of the frontal lobes. To assess and treat these patients it is important to look at their memory and executive functions. The insight that memory is more than storage has become an important part of neuropsychological science and diagnostic. Aspects like the learning skills or the use of memory strategies are increasingly from interest. Two strategies that have been studied in the context of word list learning are semantic encoding strategies, in which categorization of the presented words facilitate encoding, and story mnemonics, in which the learner tries to invent a coherent story in which the words are mentioned. While the California Verbal Learning Test, in which the subjects are required to memorize a list of 16 auditorily presented nouns from four different categories (for example fish and tools) allows for a direct evaluation of a semantic strategy, there are fewer neuropsychological data investigating the influence of a story context on verbal learning. While we know that context facilitates verbal processing in aphasic patients, there are two reasons for why frontal lobe patients might not necessarily benefit from such an encoding strategy. First, it is not clear that they are able to apply the strategy successfully, i.e., to invent a story consisting of the words to be learned. Second, if they are able to do so, the additional information might interfere with memory for the target words. The aim of this study was to investigate the differential effects of semantic categorizability and of a story context on verbal learning.

Two German verbal learning tests, only one of which contains categorizable words, were used. In both tests, the word lists are presented repeatedly. The difference between the tests is that in one test the nouns are from four semantic categories, while in the other test, no semantic categorization strategy can be applied to facilitate learning. The same words were then embedded in short stories. The stories involved these words in the same sequence as in the original test, two words in a sentence, each of them elaborated with an adjective or a prepositional phrase. Each

subject was tested in each of the four cells of the resulting 2 x 2 design, with the factors Categorizability (yes versus no), and Context (story versus word list). Which of the parallel versions appeared in which of the Context conditions, and the order of presentation of the four tests was counterbalanced across participants. Similar to the original learning procedure, each test was presented five times, so that in addition to the performance the course of learning could be evaluated.

In the first experiment, 24 healthy participants of two different age groups were tested to explore whether the two strategies would be facilitative, and, if so, whether there would be age differences in the application of the learning strategies. There was an advantage of text for both age groups. For the healthy subjects the story context aided verbal learning even if it was not self-generated but rather experimenter-provided. Only older people benefitted from the categorizabality.

In the second experiment, fifteen well-selected brain-injured patients were tested. As expected, their learning performance was lower, but as a group, the patients showed a qualitatively similar pattern as the control participants. However, in evaluating the results of individual patients we did find qualitative differences. We could distinguish two different result patterns in the no-categorization-condition. One group of patients (story+) showed a larger profit from the story context as the healthy subjects. A second group of patients (story-) had significantly lower results in the story condition as in the list condition. All the story+ patients had focal lesions in the left hemisphere and showed aphasic difficulties. These patients benefitted from the additional context in the story condition. In contrast, the story- patients had diffuse lesions after traumatic brain injury, involving the frontal lobes. They showed a very low performance level in the text condition. One possible interpretation is that the aphasic patients with focal lesions in the left hemisphere had difficulties to activate or process relevant information; therefore, providing additional, external information during encoding is facilitative. In contrast, the patients with traumatic brain injuries may show the failure to inhibit irrelevant information or to select context appropriate content, so that the additional story context interferes with the successful encoding of the relevant words. Further, we analyzed the correlation to neuropsychological tests, which assess attention, working, and long term memory, executive functions, and language abilities. From a theoretical viewpoint it is important to register that there was a different effect of semantic categorization and of embedding the nouns in a story context. From a clinical viewpoint our results may contribute to the process of making a differential diagnosis and develop and evaluate appropriate intervention.

Non-aphasic Language Deficits After Brain Injury: A Comparison of Comprehension Skills of Patients with Different Discourse Production Profiles

E. C. Ferstl*, T. Guthke**, F. Siebörger*** &

D. Y. von Cramon*/**

*Max-Planck-Institut of Cognitive Neuroscience, Leipzig (Germany)

**University of Leipzig (Germany)

***University of Osnabrück (Germany)

ferstl@cns.mpg.de

The concept of non-aphasic language deficits denotes communication deficits apparent on the pragmatic and discourse level rather than on the word and sentence level. This type of deficit is usually not measurable using standard aphasia tests, and it has been observed mostly after lesions to frontal brain regions and after traumatic brain injury (e.g., Novoa & Ardila, 1987). Descriptions of discourse production deficits include symptoms related to executive dysfunction, such as verbosity, planning deficits, sequencing problems and the failure to use context and background information. However, it has frequently been noted that patients exhibit markedly different discourse profiles (e.g., Coelho, Liles & Duffy, 1991; Hartley & Jensen, 1992).

In contrast, little is known about language comprehension of patients with non-aphasic language deficits. Few studies have investigated comprehension in traumatic brain injury (e.g., Nicholas & Brookshire, 1995), and the results are not clear-cut. While these studies show a lower level of performance overall, the qualitative performance patterns of TBI patients are comparable to those of control participants. One explanation for these results might be that patients with different "comprehension profiles" were grouped together in these studies, so that a differentiated analysis was not possible. To date, no studies have specifically targeted the question of whether and how the dissociable discourse production profiles map onto comprehension processes. Possible hypotheses linking production and comprehension deficits require the distinction of activation processes on one hand and selection processes on the other hand. Reduced speech output might reflect a failure to activate relevant semantic information, while enhanced output might reflect the failure to inhibit irrelevant information or to select context appropriate content. These processes have been ascribed to right versus left-hemispheric functions (Beeman, 1993). Alternatively, reduced speech output might be related to reduced overall drive, and thus, to strategic or goal-oriented use of language, while automatic activation processes remain largely intact.

Method

To test these hypotheses, this study used an extended single case approach. The goal was to identify patients with non-aphasic the language production that show distinct discourse profiles. Two patients were selected with reduced speech output, and two patients with enhanced speech output. For these patients, we then thoroughly compared and contrasted both discourse production and text comprehension deficits. The production tasks used to elicit discourse included a) spontaneous speech, in which the experimenter and the patient engaged in a dialogue about leisure activities; b) production of a procedural text, in which the patient was asked to describe in detail the steps required to perform a complex procedure (e.g., renovating a room); c) description of a cartoon picture sequence; d) production of a narrative, i.e., retelling of a previously heard story. These tasks differ in terms of how much information and/or structure needed to be generated by the patient. The protocols were first used to confirm the therapists' evaluation of the patients' discourse profiles by measuring quantitative variables, such as the rate of speech, the type/token ratio, or the amount of so-called maze material (repetitions, comments, interjections). In addition, the protocols were evaluated with respect to content-based and structural variables, such as coherence, informativeness, redundancy, completeness, and use of the canonical text structure.

To evaluate text comprehension processes we selected the following five paradigms: a) in a homonym priming study, we tested whether activation processes were intact; b) using a word-recognition paradigm we tested different levels of text representation; c) an off-line coherence judgment task was used to evaluate non-automatic inference processes; d) using yes/no comprehension questions we tested memory for a narrative without production requirements, and e) in a story recall experiment (see production), we evaluated retelling, as well as goal-directed retrieval processes (see Ferstl, Guthke & von Cramon, 1999). The first two paradigms enable to directly test the hypothesis of semantic activation processes playing a role for hypo- versus hyperphasic speech style. The remaining three paradigms are off-line experiments that enable to investigate intentional processes during language comprehension and language use.

Results and Discussion

The patients' performance in these tasks was compared to the performance of control groups to identify potential comprehension deficits. As expected, neither hypophasic patient showed a priming effect in the homonym priming experiment (a). Their performance in the text word recognition paradigm was roughly comparable as well (b). In contrast, the patterns of results in the off-line tasks did not correspond to the grouping according to discourse production profiles. Here we found that two patients, one hypophasic and the other hyperphasic, had difficulties with drawing strategic inferences (c). The same two patients failed to utilize the perspective cue despite high overall recall performance.

To summarize, these findings suggest that reduced speech output might in fact be related to the failure to activate relevant semantic associations. On the other hand, no such clear pattern obtained for patients with enhanced speech output in the paradigms employed here. Further research is needed to delineate possible contributions of lesion location and other neuropsychological deficits on the connection between discourse production and text comprehension.

–

References:

Beeman, M. (1993). Semantic processing in the right hemisphere may contribute to drawing inference from discourse. Brain and Language, 44, 80-120.

Coelho, C. A., Liles, B. Z., & Duffy, R. J. (1991). Discourse analysis with closed head injured adults. Evidence for differing patterns and deficits. Archives of Physical Medicine and Rehabiliation, 72, 465-468.

Ferstl, E. C., Guthke, T. & von Cramon, D. Y. (1999). Change of perspective in discourse comprehension: Encoding and retrieval processes after brain injury. Brain and Language, 70, 385-420.

Hartley, L. L., & Jensen, P. J. (1992). Three discourse profiles of closed-head-injury speakers: Theoretical and clinical implications. Brain Injury, 6, 271-282.

Nicholas, L. E. &Brookshire, R. H. (1995). Comprehension of spoken narrative discourse by adults with aphsia, right-hemisphere brain damage, or traumatic brain injury. American Journal of Speech-Language Pathology, 4, 69-81.

Novoa, O. P. & Ardila, A. (1987). Linguistic abilites in patients with prefrontal damage. Brain and Language., 30, 206-225.

Discourse Coherence and Cohesion in Aphasia

T. Djurovic

Institute for Psychophysiology and Speech Pathology (Yugoslavia)

gtdjuro@EUnet.yu

Discourse can be regarded as a multileveled object of study, since it encompasses different aspects of both linguistic and cognitive structure and organization. The production and comprehension of discourse activate, in a simultaneous fashion, varied segments such as phonology, lexicon, syntax, the concepts of cohesion and coherence and so on. Thus, one can assume that discourse being so comprehensive, the impairments in case of language pathology would also be holistic. The current research on discourse in neuropsychology emphasizes the real communication ability of brain-damaged patients, i.e. their final output in a more natural context. The testing of linguistic knowledge of aphasic patients for assessing their naming ability or production and comprehension of syntactic structures is undoubtedly invaluable, but not sufficient. This is due to the fact that language is both for healthy subjects as for aphasic subjects, a medium of not only linguistic competence, but also behaviour whose primary intention is communication. Therefore, the present day tests should take into account larger language chunks than separate words or sentences, since aphasic subjects as some tests have confirmed as well as those with intact brain and language, exhibit far too more complex patterns of communicative competence. Both healthy people and those with some language impairments communicate not by using isolated words and sentences, but rather by resorting to discourse structures, using them in a specific setting, and developing distinctive discourse styles. Moreover, the research results have revealed that aphasic patients show high degree of communicative competence, i.e. they communicate better than they speak.

Our experiment taps on the ability of aphasic subjects to overcome their deficits on word or sentence level, exhibiting a high degree of language integrity on discourse level. The aphasic patients have been tested regarding their ability to produce narrative and procedural discourse. Two crucial concepts which establish not only lexical and grammatical connectedness between sentences in a discourse, but also contribute to the overall conceptual and pragmatic unity of discourse are concepts of cohesion and coherence respectively. Consequently, the impairments of cohesion and coherence result in the dysfunctioning of different linguistic and nonlinguistic - cognitive processes. It seems that problems with cohesion as a purely linguistic phenomenon leads to impairments at microlinguistic or intrasentential level, i.e. impairments of smaller structural units at the level of single words or sentences, which function in a rather de-contextual setting. On the other hand, problems with coherence as a cognitive phenomenon result in impairments at macrolinguistic or intersentential level, since it deals with suprasyntactic discourse units, which are highly contextualized. We predicted that the concepts of cohesion and coherence would be differently spared, that is, impaired depending on the type of aphasic patients, Broca's versus Wernicke's patients. In other words, we supposed that the above named concepts are coexistent but at the same time mutually exclusive, i.e. coherence would be spared in Broca's subjects, but the elements of cohesion, both lexical and grammatical ones would be damaged. On the other hand, Wernicke's patients have been supposed to exhibit higher cohesion preservation contrary to that of coherence.

Method

Subjects. Two groups of aphasic patients have been tested, 5 subjects of Broca's type and 5 subjects of Wernicke's type. Groups were matched for age (the age range was 35-62), gender (all male subjects), education (all subjects had at least a secondary education), and for time postonset (at least six months postonset). All subjects were monolingual speakers of Serbian and were patients of the Institute for Psychophysiology and Speech Pathology in Belgrade, Yugoslavia. Etiology in all cases was vascular. For inclusion in the study, subjects were required to have received a definitive diagnosis by a neurologist (according to BDAE) and to have no visual disturbances which would hamper them to complete the required tasks.

Stimuli. Prior to elaborate research of both discourse types all aphasic subjects were required to describe a picture in order to incite discourse production. For eliciting narrative discourse, three sets of sequential pictures were used. Subjects were instructed to tell a story about each set of pictures; pictures were presented in a logical and conceptual order with the systematic appearance of participants in narrative episodes. Each set contained a subtle message which should be recognized. Procedural discourse was elicited by instructing subjects to describe two everyday situations, tooth-brushing and shopping. The subjects were suggested that the researcher did not know how to perform the requested tasks. It was estimated that the procedures were known to subjects, that is, that they could perform the task. The mentioned procedures were different regarding both linguistic and conceptual complexity. Discourse samples were audiotaped and transcribed.

Results

Although the overall structure in both narrative and procedural discourses may be preserved in aphasic patients, some level of impairments of conceptual and pragmatic discourse organization has been discerned. The severity of language impairments with Broca's subjects has prevented them to clearly describe and convey the story to the rater. The style of Broca's patients is characterized by significantly reduced cohesion elements-inconsistent use of references, almost absence of elements of deixis. Generally, Broca's patients were able to get the main gist of the story, although the information format was reduced to a large extent. As for procedural discourse, it has been noted that the main steps were included, but again the language was greatly decreased regarding its informative content. On the other hand, Wernicke's patients exhibited highly fluent speech interspersed with a lot of incongruous references, significantly higher percentage of deictic words, higher incidence of verb forms, but at the same time reduced coherence of discourse samples, both narrative and procedural. Nevertheless, the structure of discourse has been maintained with both aphasic types in spite of reduction in the amount and complexity of language. We can generally conclude that all the essential elements of the story are included with both groups of aphasic patients, they follow the chronological sequence of the events and all the participants are marked. The significant deficits could be seen in supplying the gist of a story, especially with Wernicke's patients. As for procedural discourse, all essential steps are included, although both coherence and cohesion in procedures are decreased to some extent with Broca's patients and Wernicke's patients as well. We can suggest that the impairments of discourse are in correlation with the clinical types of aphasic patients. Consequently, aphasic patients of Broca's type exhibit rather morpho-syntactic deficits, while those of Wernicke's type show more lexico-semantic deficits. They are both characterized by well-known phonological deficits such as phonemic paraphasia and/or neologisms.

The present study pursues the issue of discourse abilities in aphasic patients who speak a language quite unlike English. Therefore, we expected rather different results regarding case inflection or free-word order as clearly marked cohesive elements. Apart from unselective use of nominative case with two patients, the study demonstrated that patients were highly sensitive to inflectional violations, which we can further interpret as a preservation of this information in so-called mental, universal lexicon.

Discussion

The experiment has shown that while story grammar may be preserved with both types of aphasic patients, Broca's and Wernicke's patients, coherence is frequently lost. This made us propose that we can speak of a general mental structure of each story. Every story may have some mental structure, be it story grammar or story schema, which is inherent to every story but independent

of its content. Since our aphasic subjects have demonstrated similar language deficiencies as those speaking different languages we assume that there is some underlying story schema common to all people with some variables specific to particular language (in Serbian, incidental case errors). However, the conclusions are highly tentative and still a considerable number of both the theoretical and practical evidence is needed. Since with both groups of patients coherence was more vulnerable than cohesion (mainly marked by omission of references or deixis), we pledge for those experiments which could enhance the mutual perspectives of not only impaired subjects but their non-impaired partners as well, as to contribute to the higher degree of conceptual integrity of discourse.

"Aliens" and "Foreigners": Discourses of Exclusion : The Discourse Historical Approach

R. Wodak

University of Vienna (Austria)

Ruth.Wodak@oeaw.ac.at

In this paper, I will start of with analyzing several texts of different genres concerning the new political situation in Austria. Specifically, I will be concerned with the populist discourse of the FPÖ and similar discourses in other European countries. I will relate this analysis to a theoretical framework, provided by the Discourse-Historical approach which was developed in various studies on political discourse in the past 15 years (Reisigl & Wodak, 2000). The texts which I will analyze – inter alia -will illustrate the recent debate on immigration in various parliaments in 6 EU-countries (Wodak & van Dijk, 2000). The focus of my paper will be foremost to present this theoretical approach and the methodology used to understand discrimination of minorities. First, I will present our own Discourse Historical Approach. Politics is mainly based on discourse, thus the problem – why does the FPÖ have such success- immediately implies the investigation of the FPÖ's rhetoric in many public and private domains. My main claim in this lecture is that "context", probably the most misused term in linguistics, is THE big challenge for our discourse theories and also for our data collection and the detailed analysis. We need theories about the specific context we investigate in our research, and this is exactly the place where social theories must come in. Specifically, I would like to elaborate four levels of theorizing (Wodak, 2000, in press). Theories about populism, about coming to terms with the Nazi past ("Vergangenheitsbewältigung") and about neo liberalism and globalization have to be applied and integrated to explain some of the facets of the Haider phenomenon. Lastly, I assume that the Haider phenomenon is related to tensions between nationalism, identity and globalisation: This might be the general framework of Grand Theories. In my lecture, I see the analysis of Haider´s populism as an example for the relevance of social theory, ethnography, multiple genre analysis and detailed linguistic analysis in textlinguistics and CDA. In investigating complex social problems, we need a theoretical framework, which labels, systematizes and explains our ethnographic experiences which first form a kind of "symptomology". Our task as critical scholars is to relate relevant "symptoms" and "phenomena" with each other and to offer theoretical explanations for such relationships. Accordingly, we have to justify explicitly why we relate certain "symptoms" with each other and how we come to understand them in a certain way. This entails an hermeneutic approach in the sense of the Frankfurt School which does not mean that we do not want to explain social phenomena. The difference between "verstehen" (understand) and "erklären" (explain) is important. In my opinion, we can not aim at any kind of mono- causal explanation in the sense of the Natural Sciences. Social phenomena are much too complex and historically embedded to be explained in such uni-directional ways.

A more pragmatic approach, like the one of Nicos Moutzelis, seems adequate. In his book " Sociological Theory: What went wrong? " (Moutzelis, 1995), Moutzelis introduces the idea of "conceptual pragmatism" as a possible way out of the theory crisis in the social sciences. According to Moutzelis, social theory "has as its major task to clarify conceptual tools and to construct new ones by following criteria of utility rather than truth" (1995). Such a pragmatic approach to theory would not purport to provide a catalogue of contextless propositions and generalizations, but rather relate questions of theory formation and conceptualization closely to the specific problems that are to be investigated. In this sense, the first question we have to address as researchers is not: Do we need a Grand Theory? But rather: Which conceptual tools are relevant for this or that problem and for this and that context? Although the former question might invite exciting speculations, it leads away from problem-oriented science. Thus, I assume that the complexities of modern societies can only be grasped by a model of multicausal, mutual influences between different groups of persons within a specific society.

In investigating historical, organizational and political topics and texts, the discourse-historical approach attempts to integrate much available knowledge about the historical sources and the background of the social and political fields in which discursive "events" are embedded. Further, it analyzes the historical dimension of discursive actions by exploring the ways in which particular genres of discourse are subject to diachronic change (Wodak et al., 1990, 1994; Wodak 1996). Lastly, and most importantly, this is not only viewed as "information": At this point we integrate social theories to be able to explain the so-called context. In the following, I would like to give some short definitions which I will illustrate in my lecture by analyzing parliamentary debates on immigration in EU countries.

Discourse. In accordance to other approaches devoted to Critical Discourse Analysis, the discourse-historical approach perceives both written and spoken language as a form of social practice (Fairclough & Wodak, 1997). A discourse is a way of signifying a particular domain of social practice from a particular perspective (Fairclough, 1995). As critical discourse analysts, we assume a dialectical relationship between particular discursive practices and the specific fields of action (including situations, institutional frames and social structures) in which they are embedded. On the one hand, the situational, institutional and social settings shape and affect discourses, and on the other, discourses influence discursive as well as non-discursive social and political processes and actions. In other words, discourses as linguistic social practices can be seen as constituting non-discursive and discursive social practices and, at the same time, as being constituted by them. To put it more precisely: "discourse" can be understood as a complex bundle of simultaneous and sequential interrelated linguistic acts. They manifest themselves within and across the social fields of action as thematically interrelated semiotic, oral or written tokens, very often as "texts", that belong to specific semiotic types, i.e. genres.

Text. We conceive "texts" as materially durable products of linguistic actions, as communicatively dissociated, "dilated" linguistic actions that during their reception are disembedded from their situation of production (Ehrlich, 1983; Graefen, 1997; Reisigl, 2000, in press).

Genre. A "genre" may be characterized, following Norman Fairclough, as the conventionalized, more or less schematically fixed use of language associated with a particular activity, as "a socially ratified way of using language in connection with a particular type of social activity" (Fairclough, 1995).

Fields of Action. "Fields of action" (Girnth, 1996) may be understood as segments of the respective societal "reality", which contribute to constituting and shaping the "frame" of discourse. The spatio-metaphorical distinction among different fields of action can be understood as a distinction among different functions or socially institutionalized aims of discursive practices. Thus in the area of political action we distinguish between the functions of legislation, self-presentation, the manufacturing of public-opinion, developing party-internal consent, advertising and vote-getting, governing as well as executing, and controlling as well as expressing (oppositional) dissent. We can

represent the relationship between fields of action, genres and discourse topics with the example of the area of political action.

Field of action 1: Law-Making Procedure; Field of action 2: Formation of Public Opinion & Self-Presentation; Field of action 3: Party-internal development of an informed opinion; Field of action 4: Political advertising/propaganda; Field of action 5: Political executive/Administration; Field of action 6: Political control.

A "discourse" about a specific topic can find its starting point within one field of action and proceed through another one. Discourses and discourse topics "spread" to different fields and discourses. They cross between fields, overlap, refer to each other or are in some other way socio-functionally linked with each other.

Interdiscursivity and Intertextuality: Interdiscursive (e.g. the intersection between two discourses) and intertextual relationships between discourses - whether of an explicitly referential kind, a formally or structurally iconic (diagrammatical) kind, or in the form of tropological correlations, evocations, allusions or (direct and indirect) quotations – can be represented schematically. Also, the assignment of texts to genres, the topics to which a text refers, the topical intersection of different texts can be easily displayed. Finally, the specific intertextual relationship of thematic reference of one text to another can also be represented.

Strategies: (1) How are persons named and referred to linguistically? (2) What traits, characteristics, qualities and features are attributed to them? (3) By means of what arguments and argumentation schemes do specific persons or social groups try to justify and legitimize the exclusion, discrimination, suppression and exploitation of others? (4) From what perspective or point of view are these namings, attributions and arguments expressed? (5) Are the respective discriminating utterances articulated overtly, are they even intensified or are they mitigated?

By "strategy" we generally mean a more or less accurate and more or less intentional plan of practices (including discursive practices) adopted to achieve a particular social, political, psychological or linguistic aim. As far as the discursive strategies are concerned, that is to say, systematic ways of using language, we locate them at different levels of linguistic organization and complexity.