DISCOURSE CONNECTIVES AND LEXICAL COHESION: AN EXPERIMENTAL INVESTIGATION OF BI-CLAUSAL SENTENCE PROCESSING IN TURKISH

Gökhan Gönül (gokhan.gonul@metu.edu.tr) **Deniz Zeyrek (dezeyrek@metu.edu.tr)**

Cognitive Science, Informatics Institute
Middle East Technical University, 06800, Ankara, Turkey

Abstract

It is a widely accepted fact that coherence enables a text's comprehensibility. A major source of coherence is discourse cohesion (textual properties of the text). Lexical cohesion (e.g. synonymy) and discourse connectives are two major types of discourse cohesion. We investigate the contribution of these two types of cohesion to the overall comprehension of bi-clausal sentences in Turkish. In a two-phase study, we ask the participants to judge the comprehensibility of sentences while we obtain eye-gaze data and then ask them to write recall protocols. We find that lexically cohesive sentences (labeled as high coherent) are judged more comprehensible and recalled better, and that in low coherent sentences (those lacking lexical cohesion), the fixation counts are high. This study shows that in short texts, lexical cohesion guides coherence and it is singled out as an important factor of discourse comprehension. The study concerns Turkish discourse and may have implications on discourse coherence and discourse comprehension in other languages.

Keywords: coherence, lexical cohesion, contrastive discourse connectives, comprehension, eye-tracking, reading.

Introduction

People read a great deal of written material every day, including newspapers, textbooks, research papers, the material on the internet, etc. For the reader, the comprehension of written (or spoken) material is simply a reflexive behavior, but it is of interest to cognitive scientists, linguists and psycholinguists to understand what exactly causes the comprehensibility of a stretch of discourse. All approaches to discourse maintain that discourse comprehension is a cognitive process that arises from the various sources of knowledge accessible to the reader.

Comprehension of texts is an integrated process, which includes making sense of the individual sentences and forming the gist of the whole text. In their construction-integration (CI) model of text comprehension, Kintsch & van Dijk (1978) propose that these levels are integrated by the reader. These levels involve forming the text base, where the sentences are parsed and the meaning of individual sentences is constructed; and the level where a global text understanding is formed by integrating only those propositions from the text base that fit well with the context of the text (also see McNamara, 2001; Kintsch, 1994). *Construction* (of meaning) is based on textual properties and propositions conveyed in the text. *Integrations* are processes in the readers' cognition activated with the help of the text; this activation is a dynamic process which continues throughout the text.

The text-based mental representations are the propositional networks that are created by the text and are developed in the *construction stage*. The text has many surface properties helping the construction process, e.g. sentence connectives, lexical ties, pronouns. They connect the prior discourse context to the current discourse context, and help text-based memory or understanding. The text-based representations are affected by propositional representations of texts and the readers' local inferences. The whole process of construction and integration creates a coherent text for a reader or listener.

In the current understanding of discourse (this is our term for 'text', regardless of whether it is written or spoken), the role of the reader is essential; to make sense of a piece of text, the reader brings together parts of discourse to form interpretative structures (Cornish, 2009, Halliday, 1994). Discourse makes sense thanks to both discourse coherence (global interpretative structures) and discourse cohesion (lexical ties, pronouns, ellipsis, etc.). In this paper, we are particularly interested in discourse coherence to the extent it is guided by discourse cohesion, namely discourse connectives and lexical links in sentences. Drawing inferences from a span of discourse (e.g. via presuppositions or conversational implicatures) are out of the scope of this study. In the construction-integration model, cohesive elements are helpful at the construction level. In this study, we propose and test the assumption that in discourse (such as bi-clausal sentences) two types of cohesive elements, namely lexical relations (in our case, synonymous ties) and connectives are competing factors in discourse processing. This assumption has not been tested before in Turkish.

In this paper, we take discourse connectives as lexical anchors making a discourse relation explicit, such as *Contrast, Cause-Effect, Expansion, Temporal*, etc. (Kehler 2002, Mann & Thompson 1986, PDTB Research Group, 2008). It is known that discourse connectives exist in all languages and make important contributions to discourse coherence (Knott, Sanders et al., Mann & Thompson, 1986). They can be drawn from natural language conjunctions, e.g. coordinating conjunctions (*and, but*), subordinating conjunctions (*because*), and discourse adverbials (*however, therefore*). We will say that these connectors act as discourse connectives when they relate two clauses. For example, in (a), *and* is a discourse connective, while in (b),

it is not, because it links two nouns rather than clauses. Connectives like (b) are out of scope of the present paper.

a. Jane went to the store and Tim remained at home.
b. Jane and Tim went to the store.

The second type of cohesive device we deal with, i.e. lexical cohesive devices involve repetition (the repetition of the same word), synonymy (words with same or very similar meanings), and collocation (defined as a co-occurrence tendency) (Halliday, 1994). While we keep repetition and collocation out of the scope of this paper, we concern ourselves with the category synonymy, which includes: synonymy, antonymy (words with opposite meanings), hyponymy (a relation of inclusion) and meronymy (the part/whole relationship). Saeed (2003:65-70) provides the following examples for these lexical ties: couch/sofa (synonyms), death/life (antonyms), dog/animal (hyponym), page/book (meronym).

Following early works on discourse comprehension (e.g. McNamara, et al. 1996), we take it a fact that coherence affects a text's comprehensibility. To the best of our knowledge, studies investigating the role of discourse cohesion on discourse comprehension are few but they exist. For example, Millis and Just (1994) report that subjects recognize the verb from the first conjunct faster when the statements are conjoined with connectives (*because, although)*, and read the upcoming sentence more quickly. In a recent eye-tracking study, Köhne & Demberg (2013) find that concessive discourse markers in German (*Dennoch* 'nevertheless') can be processed rapidly if the visual context is constraining enough, while causal connectives (*Daher* 'therefore') are processed with a delay.

Regarding naturally occurring language, it is clear that discourse may be regarded as perfectly coherent even though it lacks a discourse connective, suggesting that other sources establish coherence in cases where a discourse connective is lacking. In this paper, we aim to tackle this issue experimentally, limiting ourselves with two contrastive discourse connectives in Turkish (*ama* 'but, yet', *fakat* 'but') and lexical cohesion (synonymy). Firstly, we wanted to see whether coherence arising from these two sources facilitates comprehension in Turkish as in the studies with other languages. Secondly, we asked whether lexical cohesion might be a more powerful cue than discourse connectives for discourse comprehension.

We tested our predictions via an with on-line (eye-tracking) and an off-line task (i.e., comprehensibility judgments and recall protocols). Based on our foci in the study, we made the following predictions. In the rest of the study, the sentences which have lexical cohesion are referred to as high coherent sentences (HCoh) and the ones without lexical cohesion are named as low coherent (LCoh) sentences.

- Comprehensibility judgment results will be higher for HCoh sentences than the LCoh sentences. Moreover, comprehensibility judgment scores will exhibit a worsening comprehensibility trend from HCoh sentences (G1, G2, see below) to LCoh sentences (G3, G4, see below).
- The HCoh sentences will be better recalled than the LCoh sentences and the recall results will decrease steadily from from HCoh sentences (G1, G2) to LCoh sentences (G3, G4).
- Fixation counts per one word will be lower for the HCoh sentences (G1, G2) than the LCoh sentences (G3, G4). Moreover, the fixation counts per one word will increase steadily from HCoh sentences (G1, G2) to LCoh sentences (G3, G4).
- HCoh sentences with no discourse connective (G2) will be recalled and processed similarly to HCoh sentences with a discourse connective (G1); overall, HCoh sentences (G1 & G2) will be recalled and processed better than LCoh sentences (G3 and G4).

What we take as high coherent and low coherent texts are explained and illustrated in Table 1. In the examples, the discourse connective is underlined; lexical ties are shown in italics.

Method

Participants

We tested 46 right-handed native Turkish speakers with healthy eyes, 3 of whom withdrew from the experiments explaining that they had a breathing problem related to articulatory suppression (explained below). In addition, the data from 3 participants were not analyzed since their data were substantially lower than that of the other participants; upon examination of their data, it appeared that they were unable to understand the procedure. Data from 40 participants (27 females 13 males) were analyzed. Only 1 participant's data were not analyzed with respect to eye movements because of the problems with the eye-tracker. The age of the participants varied from 21 to 37 (M=23, S.D=3.93). The participants were randomly placed in 4 experimental groups so that all the groups had 10 participants.

Materials, Design and Procedure

We constructed four groups of stimuli which were bi-clausal sentences having or lacking lexical cohesion with or without a contrastive discourse connective. All the words in the stimuli were checked for frequency of use (Göz, 2003), and only the high frequency words were included in the stimuli. Two linguists checked all the test sentences, ensuring that they were grammatical and coherent (i.e., made sense). They also confirmed that G1 and G2 sentences have lexical cohesion while G3 and G4 sentences lacked lexical cohesion.

Group 1 (G1) sentences constituted lexically cohesive clauses and a contrastive discourse connective, *ama* 'yet, but' or *fakat* 'but', Group 2 (G2) sentences had lexically cohesive clauses with no discourse connective, Group 3 (G3) sentences had disrupted lexical cohesion but contained the discourse connective, Group (G4) sentences had

disrupted lexical cohesion and lacked the discourse connective. G1 and G2 were labeled the high coherent group (HCoh), G3 and G4 were the low coherent group (LCoh) (Table 1). Lexical cohesion was ensured by synonymy, meronymy, antonymy, and hyponymy. For example, in Table 1, G1 and G2 have antonymic lexical cohesion relation (*death* in the first clause and *life* in the second clause), however the highlighted words in G3 and G4 (*death* in the first clause, *nature* in the second clause) do not present clear antonymy relations. The total number of words in the target sentences differed between minimum 4, maximum 13 (G1: min=4, max=13, total=111; G2: min=4, max=12, total=99; G3: min 5, max=12, total=105; G4: min=4, max=11, total=93).

The sentences were presented to the participants in size 15 Times New Roman fonts in black color on a white screen. All the participants saw 24 sentences randomly (12 distractors, 12 target sentences) after a training session with 6 sentences.

The experimental procedure consisted of two consequently ordered phases; i.e., reading the sentence and undertaking a comprehensibility judgment test, and typing a recall protocol of the sentence.

Table 1: Examples of the stimuli used in the experiment (Lexcoh: lexical cohesion; DC: Discourse connective)

	Lexcoh	DC
(G1) Bir tabloda *ölümü* anlatmak kolaydır ama *yaşamı* resmetmek emek ister. 'It is easy to *convey death* in painting but *to portray life* requires effort'.	Yes	Yes
(G2) Bir tabloda *ölümü* anlatmak kolaydır, *yaşamı* resmetmek emek ister. 'It is easy to *convey death* in painting, to *portray life* requires effort'.	Yes	No
(G3) Bir tabloda *ölümü* anlatmak kolaydır ama *doğayı* resmetmek emek ister. 'It is easy to convey *death* in painting but to portray *nature* requires effort'.	No	Yes
(G4) Bir tabloda *ölümü* anlatmak kolaydır, *doğayı* resmetmek emek ister. 'It is easy to convey *death* in painting, to portray *nature* requires effort'.	No	No

Phase I Throughout the whole experiment, participants' eye-gazes were traced by an eye-tracking device (explained below). The participants' eye movements were calibrated before the experiment. The experiment started when the participants saw a fixation point (for 2000 ms). As they started to read the sentences, they were asked to articulate the sound [b] intermittently. This is the articulatory suppression technique, ensuring that the participant's ability to use auditory cues to encode information in the (working) memory is inhibited (Baddeley, 1992; Larsen and Baddeley, 2003). Kintsch & van Dijk (1978) emphasize that coherence representations were under the limitations of working memory. We know from the memory literature that people may remember well-structured sentences easily (Jefferies, Ralph & Baddaley, 2004). In long texts, there are various elements for loading the working memory, forcing the reader to make inferences from the text (i.e. at the integration level in the CI model). In a similar way, given that lexical cohesion is effective at the intra-sentential level or among close sentences in texts, it was not implausible for our subjects to memorize the words in the test sentences, which are not long texts. Therefore, we aimed to create a working memory load with the articulatory suppression technique so as to obtain reliable results regarding the effect of meaning construction in short texts.

The participants' voices were recorded to ensure all participants articulated the [b] sound. (None of the participants whose data were analyzed failed to articulate the [b] sound during the experiment). The participants were asked not to move their head while they were reading the sentences. They clicked the 'enter' button as soon as they understood the sentence (and they stopped repeating the sound [b]), and they were presented with a comprehensibility judgment question. This question asks the participant to evaluate the comprehensibility of the sentence on a scale of 1 to 6, where '1' means 'totally comprehensible' and '6' means ' totally incomprehensible'.

Phase II After recording their comprehensibility judgment score, the participants were asked to type a recall protocol of the sentence they have just read and evaluated in terms of comprehensibility. Only after the target sentences they were asked to answer a judgment question and type a recall protocol. Half of the distractors were bi-clausal sentences with or without a DC.

The participants' eye movements were traced by the Tobii Studio T-120 eye-tracker, where the data rate is 120Hz and accuracy is 0.5 degrees. The spatial resolution of the eye-tracker is 0.3 degrees with 0.1 degrees for drift and 0.2 degrees for head movement error. The latency is maximum 33 ms, and the blink tracking recovery takes maximum 33 ms. The time to tracking recovery is 300ms.

Data Analysis and Results

Analyses were conducted with the variables Coherence (HCoh & LCoh) and Group (G1, G2, G3, G4) with respect to the comprehensibility judgment task and the recall results. The comprehensibility judgment scores were quantified by taking the sum of all the judgments in

respective groups of sentences. The recall was calculated per word. The total number of recalled discourse connectives and the total number of discourse connectives were removed from the analysis in order to compare groups. The following formula was used:

[(sum of words in target sentences) – (sum of discourse connectives in target sentences)]

[(sum of recalled words in target sentences) – (sum of recalled discourse connectives in target sentences)]

Though the participants' eye movements throughout the whole procedure were traced, only the data obtained while they were reading the target sentences were analyzed (i.e. the data from their first fixation to their last fixation before answering the comprehensibility question).

The number of single fixations in the sentences was counted and the result was divided by the number of total words in sentences. This is called fixation count per word and was our independent variable. Only the fixation count data were used.

Effect of coherence on comprehension

The t-test results of recall showed that there was a significant effect of coherence, t(39)=2.71, p=.01; the words in the HCoh group (M=0.85; S.D.=0.06) were better recalled than those in the LCoh group (M=0.75; S.D=0.13). The results are represented in a graph in Figure 1. We used an alpha level of .05 for all statistical tests.

On the other hand, there was no significant effect of coherence on comprehensibility judgments, t(39)=-1.74, p=.09.

The mean of the fixation count of per word was significant between the HCoh and LCoh groups, t(38)=-2.265, p=.029; the mean of fixation counts per word was lower in the HCoh groups (M=1.97, S.D= 0.49) than in the LCoh groups (M=2.32, S.D=0.48). These results can be seen in graphically in Figure 2.

The differential effect of lexical cohesion and discourse connective on comprehension

In order to see the differential effect of lexical cohesion and discourse connective on comprehension, we analyzed group differences. We analyzed the following pairs of sentences with planned contrast tests.

- LexCoh with or without DC (G1 x G2)
- LexCoh and DC (G2 x G3)
- LexCoh (G1 with G2 x G4 and G1 x G3),
- LexCoh with DC (G1 x G4)

In what follows, only significant results are reported.

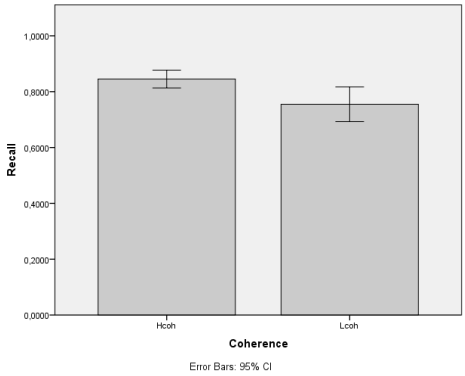

Figure1: Means of recall per word for high cohesive or low cohesive clauses

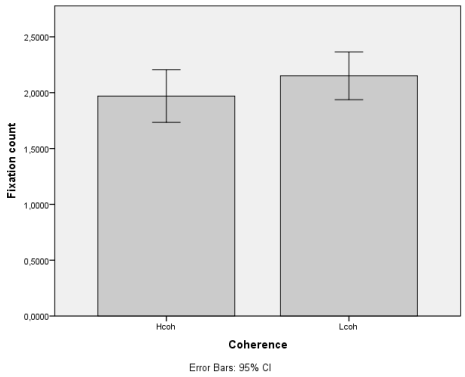

Figure 2: Means of fixation counts per word for high and low cohesion groups

For the recall test, one way ANOVA results showed that there was a significant relation between groups, F(1,39)=1.924, p=.04. The analyses revealed a significant relation between the LexCoh groups (G1 and G2) and G3 (cohesion by a DC), t(39)=2.771 (two-tailed), p=.009. The analyses also indicated a significant relation between G1 and G3, t(39)=2.75 (two-tailed), p=.009. The relation between G1 and G4 was close to the significance level, t(39)=1.979 (two-tailed), p=.055, though not reaching significance (Figure 3).

The comprehensibility judgment scores exhibited a gradual worsening from G1 to G4 and they differed significantly, F(3,36)=3.01, p=.042. The relation between G1 (M=16.11; SD=4.45) and G4 (M=22.8; SD=7.15), t (39) = -2.08, p<.045 was significant.

The group results showed that the mean of fixation counts per word increased gradually from G1 (M=1.93, SD=0.54), to G2 (M=2, SD=0.47), G3 (M=2.17, SD=0.32), and G4 (M=2.51, SD=0.6), F (1, 38)=6.67, p=. 012 (Figure 4).

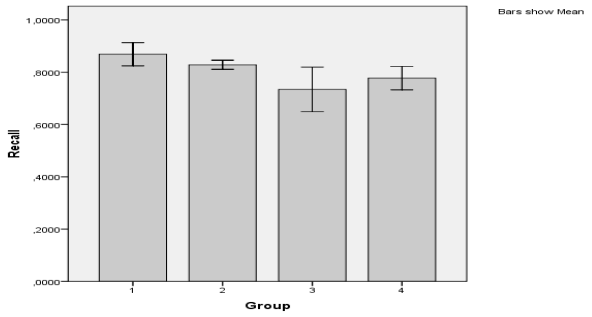

Figure 3: Means of recall per word for groups

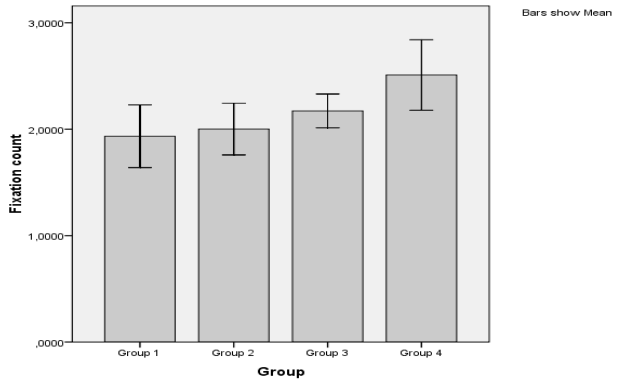

Figure 4: Means of fixation counts per word showing the differences between groups

Discussion and Conclusion

This study predicted that coherence arising from lexical cohesion and discourse connectives would increase comprehensibility (McNamara et al., 1996, McNamara and Kintsch, 1996, and McNamara, 2001). We took coherence as a matter of well-connected texts, which involved either a discourse connective and lexical cohesion, or only lexical cohesion. We ran a two-phase experiment where we tested native speakers' judgments about the comprehensibility of bi-clausal sentences, their recall protocols of the stimuli, and obtained eye-tracking data while they were reading the stimuli. Recall results and comprehensibility judgment results showed that high coherent sentences were recalled better and gave better comprehensibility judgments, showing the facilitative role of coherence both with online and offline methods using Turkish bi-clausal sentences. Furthermore, we found that the comprehensibility judgments are aligned with the results from recall protocols. The fixation count results showed that low coherent sentences had more fixation counts per word than high coherent ones. The sentences with lexical cohesion and a (contrastive) discourse connective (G1) and those with only lexical cohesion (G2) were recalled and evaluated better than the sentences which lacked lexical cohesion and had a contrastive DC (G3) and sentences which lacked both kinds of cohesion (G4). Within the perspective of construction-integration theory, we can say that lexical ties, which provide lexical cohesion in the sentences, are effective both in construction and integration stage of sentence comprehension. These results suggest that lexical cohesion has a stronger effect on coherence (and hence comprehensibility) of bi-clausal sentences. In this way, we have shown that among the two major elements of discourse cohesion, i.e. lexical cohesion (synonymy in our case) and (contrastive) discourse connectives, it is lexical cohesion that has a stronger effect on coherence and comprehensibility. To the best of our knowledge, this has not been shown before in Turkish.

The eye movement data in our study show that in less coherent texts, the fixation counts are high. Our results are parallel with other eye movement studies, for example, Carroll and Slowiaczek (1986) point to the facilitative role of lexically associative words in sentences because the patterns of eye movements change according to the effects of words or phrases.

From a discourse processing perspective, an important question in our study was whether lexical cohesion would suffice for coherence; i.e. we wanted to see whether (in bi-clausal sentences), lexical cohesion without the linking role of a discourse connective would change the native speakers' recall and processing. The recall results from HCoh sentences (G1 and G2) suggested that lexical cohesion is quite adequate for readers to derive a meaning of the text even when the sentences lacked a discourse connective. Additionally, the between-group differences in the recall experiment showed that the relation between G1 and G2 is not significant. However, the relation between G1 with G2 and G3 is highly significant, with a clear bias for G1 with G2. Similarly, eye movement data revealed that the sentences which had lexical relations (G1 and G2) were processed more easily than the ones which have disrupted lexical relations. We interpret this result as evidence for the facilitative role of lexical cohesion in guiding coherence (and comprehensibility). In short, all these results point out the fact that lexically cohesive clauses suffice for interpreting a sentence as coherent, but the reverse is not true; i.e., discourse connectives (in this experiment, contrastive discourse connectives) alone are not sufficient to interpret a sentence as coherent. This result that we found for Turkish may have implications for discourse coherence and discourse comprehension for other languages.

To conclude, although we found answers to the research questions we asked, the current study is not without limitations. For example, the articulatory suppression technique was used in order to create a cognitive load for the phonological loop and make difficult to recall items (Schendel & Palmer, 2007). We did not control participants with a high and low memory span. It is possible that the high memory span participants are affected less from the articulatory suppression technique (Baddaley, 2003). Another issue is that we only tested sentences with a contrastive discourse connective. Although contrastive discourse relations is one of the most commonly occurring discourse relations in Turkish (Zeyrek, to appear), we aim to

test the interaction of lexical cohesion with other discourse connectives in the future. Finally, further studies are needed to understand whether the effect we found for lexical cohesion in bi-clausal sentences exists in longer texts.

Acknowledgments

We thank Cengiz Acartürk for his help in measuring the eye-gaze data, Ceyhan Temürcü for his help in constructing the test material.

References

Birner, B. (2006). Inferential relations and noncanonical word order. In B. J. & Ward, G. (2006). Drawing the boundaries of meaning. 31-53.

Baddeley, A. (1992). Working memory. *Science*, 255(5044), 556-559.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829-839.

Carroll, P., & Slowiaczek, M. L. (1986). Constraints on semantic priming in reading: A fixation time analysis. *Memory & Cognition*, 14(6), 509-522.

Cornish, F. (2009). Inter-sentential anaphora and coherence relations in discourse: a perfect match. *Language and Sciences*, 31: 572-592.

Göz, İ. (2003). *Yazılı Türkçenin Kelime Sıklığı Sözlüğü [Word Frequencies in Written Turkish]*. Türk Dil Kurumu Yayınları.

Köhne, J. & Demberg, V. (2013). . In Knauff, M., Pauen, M., Sebanz, N. & Wachsmuth, I. (Eds.) *Proceedings of the 35th Annual Conference of the Cognitive Science Society (CogSci 2013, Berlin)*. Austin TX: Cognitive Science Society.

Halliday, M. A. K. (1994). *An introduction to functional grammar*. (2nd ed.). London.

Jefferies, E., Lambon Ralph, M. A., & Baddeley, A. D. (2004). Automatic and controlled processing in sentence recall: The role of long-term and working memory. *Journal of Memory and Language*, 51(4), 623-643.

Kehler, Andrew. 2002. *Coherence in discourse*. Center for the Study of Language and Information: Stanford.

Kintsch, W. (1988). The use of knowledge in discourse processing: A construction-integration model. Psychological Review, 95, 163–182.

Knott, A. & Sanders, T. (1998). The classification of coherence relations and their linguistic markers: an exploration of two languages. *Journal of Pragmatics* 30:135-175 (1998).

Larsen, J. D., & Baddeley, A. (2003). Disruption of verbal STM by irrelevant speech, articulatory suppression, and manual tapping: Do they have a common source?. *Quarterly Journal of Experimental Psychology Section A*, 56(8), 1249-1268.

Mann, W. C., & Thompson, S. A. (1986). Relational propositions in discourse. *Discourse processes*, 9(1), 57-90.

Millis, K. K., & Just, M. A. (1994). The influence of connectives on sentence comprehension. *Journal of Memory and Language*, 33(1), 128-147.

McNamara, D. S., & Kintsch, W. (1996). Learning from texts: Effects of prior knowledge and text coherence. *Discourse processes*, 22(3), 247-288.

McNamara, D. S., Kintsch, E., Songer, N. B., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and instruction*,14(1), 1-43.

McNamara, D. S. (2001). Reading both high-coherence and low-coherence texts: Effects of text sequence and prior knowledge. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie exparimentale,* 55.1: 51.

Reitbauer, M. (2008). Keep an eye on information processing: Eye tracking evidence for the influence of hypertext structures on navigational behavior and textual complexity. *LSP and professional communication* (2001-2008), 8(2).

Saeed, John I. 2003. *Semantics (Introducing Linguistics)*. 2nd edition. Blackwell.

Schendel, Z. A., & Palmer, C. (2007). Suppression effects on musical and verbal memory. *Memory & cognition,* 35(4), 640-650.

The PDTB Research Group (2008). The Penn Discourse TreeBank 2.0 Annotation Manual. Available at http://www.seas.upenn.edu/˜pdtb/, or as part of the download of LDC2008T05.

Zeyrek, D. (to appear). On the distribution of the contrastive-concessive discourse connectives *ama* 'but/yet' and *fakat* 'but' in written Turkish. *Languages of Europe and Northern and Central Asia.* John Benjamins Publishing Company.