

Time Course of Metaphor Comprehension in the Visual World

Seana Coulson (scoulson@ucsd.edu)

Department of Cognitive Science (0515), 9500 Gilman Drive
La Jolla, CA 92093 USA

Tristan Davenport (trdavenp@cogsci.ucsd.edu)

Department of Linguistics, 9500 Gilman Drive
La Jolla, CA 92093 USA

Pia Knoeferle (knoeferl@cit-ec.uni-bielefeld.de)

Cognitive Interaction Technology (CITEC), Room 2.036
D-33615 Bielefeld GERMANY

Sarah Creel (screel@ucsd.edu)

Department of Cognitive Science (0515), 9500 Gilman Drive
La Jolla, CA 92093 USA

Abstract

To investigate the real time processing of metaphoric adjectives, we measured participants looking behavior as they listened to sentences such as *The little boy was shocked as a result of the [electrical socket/report card]* in the context of a display with four images. Displays included two Unrelated pictures, a Literal picture consistent with the literal interpretation of the adjective (an electrical socket), and a Metaphor picture consistent with the metaphorical interpretation (a report card). Sentences were divided into those with a preferred literal versus metaphorical reading of the adjective based on a norming study involving sentence fragments without the disambiguating information. Although conducted with different participants, those preferences were predictive of looking behavior during the eye tracking study. During the 1s interval before the onset of the disambiguating word, participants were more likely to fixate the image consistent with the preferred interpretation of the adjective than the unrelated pictures. That is, they were more likely to fixate the Literal picture in Literal biased sentences, and the Metaphor picture in Metaphor biased sentences. After the disambiguating information, participants showed an increased probability to fixate the actual target item, regardless of the preferred reading of the adjective. Results argue against models of metaphor comprehension that posit parallel activation of literal and metaphoric meaning.

Keywords: eye tracking, figurative language, language comprehension, metaphor, nonliteral meaning, visual world paradigm

Introduction

A central question in research on metaphor comprehension is whether and when literal meanings are activated in the course of understanding metaphors. One long-standing view is that metaphor comprehension differs fundamentally from the processing of literal language, and that the computation of a metaphoric interpretation occurs only after accessing the literal meaning, and rejecting it (e.g., Grice, 1976).

Existing studies, however, have provided little support for this position. In a classic study, Glucksberg and colleagues found that it took people longer to judge that metaphoric statements such as *Some jobs are jails* were literally false than to reject scrambled metaphors such as *Some jails are jobs*, as if they were obligatorily accessing the sensible metaphoric interpretation before the literal interpretation required by the task (Glucksberg, Gildea, & Bookin, 1982). Reading time studies have further revealed that, at least when metaphors occurred in appropriate contexts, they were read as fast as literal expressions (e.g., Gibbs, et al., 1997; see also Frisson & Pickering, 1999 for related evidence on metonymy), a finding that further undermines literal first models of metaphor comprehension.

Parallel Activation

Another prominent approach to metaphor comprehension is motivated by work in cognitive linguistics that suggests metaphoric meanings arise via analogical extension from literal ones (Coulson & Oakley, 2005; Lakoff & Johnson, 1999). Such approaches suggest the activation of concepts from the source domain related to the literal meaning play a key role in establishing the meaning of a metaphor (see, e.g., Coulson & Matlock, 2001). Although not processing models, these approaches imply that aspects of both a given word's literal and metaphorical meanings might be activated in parallel in the course of real time language processing.

A popular alternative, the Career of Metaphor, proposes that the analogical computation of metaphoric meaning posited by cognitive linguists occurs exclusively for novel metaphors; whereas the meanings of conventional metaphors involve a categorization process (Gentner & Bowdle, 2002). As most metaphoric language is highly conventional (Steen, 2008), one might expect the comprehension of metaphor to be affected by factors known

to operate more generally in language comprehension, such as meaning frequencies and contextual biases (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994). This prediction is tested in the present study.

Metaphor Comprehension in the Visual World

While most existing studies have examined how metaphors are processed in linguistic contexts, comparatively little is known about metaphor processing in relevant visual contexts, and how metaphor comprehension interacts with visual processes. A notable exception is an eye-tracking study by Richardson and Matlock (2007). In this study, participants inspected scenes showing paths (roads, rivers, and pipelines) while they listened to sentences involving fictive motion (e.g., *The road goes through the desert*) or literal controls (e.g., *The road is in the desert*). They found that with fictive motion sentences, participants spent more time inspecting the path in the image relative to control sentences. Inspection time and eye movements on the path increased even more when the terrain was first described as difficult relative to when it was described as easy to traverse. Richardson and Matlock (2007) interpreted this finding as evidence for the view that figurative language “can have an immediate effect on how we look at the world” (2007, p. 136), because figurative language activates spatial representations while the corresponding literal descriptions do not.

The study by Richardson and Matlock (2007) made a first important step in examining the interaction of visual processes and figurative language. However, the measures that Richardson and Matlock (2007) reported (total looking time and total number of eye movements to a scene region over the entire sentence) were non-incremental, thus leaving open questions with respect to the time course of figurative language processing. Crucially, their study did not address whether literal or metaphoric meanings are accessed first, or whether both of these meanings are accessed in parallel.

Moreover, extant work on whether literal and metaphoric meanings are processed serially or in parallel has focused almost exclusively on nouns. The present study will address the activation of literal versus metaphoric meanings of adjectives in sentences using the visual world paradigm (monitoring a listener’s eye movements as s/he listens to sentences that describe part of a visual scene, as in Altmann & Kamide, 1999; Tanenhaus et al., 1995). This method allows us to analyze eye movements time-locked to individual words in a spoken sentence, thus providing a window into the word-by-word processing of nonliteral language in visual contexts.

Experimental sentences thus involved polysemous adjectives with both literal and metaphorical readings, and the activation of the different meanings was indexed by the likelihood that participants fixated images compatible with each of the meanings of the critical word. Moreover, to determine the extent to which contextual factors modulate the activation of literal versus metaphorical meanings,

sentences were divided into those that had a preferred literal reading of the adjective, and those that had a preferred metaphorical reading.

Methods

Participants

Participants were 38 UC San Diego undergraduates who received experiment participation credit for a cognitive science, linguistics, or psychology course.

Materials

Experimental sentences were developed from 24 sentence templates, each with 4 variations to produce a total of 96 sentences. Experimental items described an agent (e.g. “the little boy”) with one of three types of adjectives, an ambiguous adjective (e.g. “shocked”), compatible both with a literal and a metaphoric reading, an unambiguous literal adjective (e.g. “electrocuted”), and an unambiguous abstract adjective related to the metaphoric interpretation (e.g., “surprised”). Critical sentences each concluded with a noun phrase implying either the literal, concrete meaning of the adjective (“as a result of the electrical socket,”) or the abstract, metaphoric meaning (“as a result of the report card,”). See table 1 for an example.

Four stimulus lists were constructed, each containing 24 experimental sentences (6 in each of the 4 conditions), so that no individual participant heard more than one variant of each experimental item. Each list also contained 48 filler sentences. Filler items were more diverse, and contained a variety of simple syntactic constructions (e.g. “The judge had a powder-white wig,”).

Experimental sentences were each presented with 4 clip art images, 2 targets, one designed to be congruent with both the ambiguous literal and unambiguous literal sentences, and one designed to be congruent with both the ambiguous metaphor and unambiguous metaphor sentences, along with 2 distractor images. Distractor images were taken as a pair from a different target item, and thus involved one image that was congruent with the literal reading of its actual target, and one with the metaphoric reading. Each image was thus presented twice, once as a target, and once as a distractor. Further, while the literal and metaphor images always appeared together as a pair of target items and as a pair of distractor items, all four items appeared together only once.

Table 1. Sample Quartet of Experimental Items.

Ambig. Literal	The little boy was	shocked	as a result of the	electrical socket.
Ambig. Metaphor	The little boy was	shocked	as a result of the	report card.

Unambig. Literal	The little boy was	electrocuted	as a result of the	electrical socket
Unambig. Metaphor	The little boy was	surprised	as a result of the	report card.

Figure 1a, for example, shows the target image for the Ambiguous and Unambiguous Literal sentences in Table 1 (“The little boy was shocked/electrocuted as a result of the electrical socket.”), while Figure 1b shows the target image for the Ambiguous and Unambiguous metaphor versions (“The little boy was shocked/surprised as a result of the report card.”). The relationship between sentences and the visual images often required the listener to make inferential connections, and this was the case for both literal and metaphoric images.

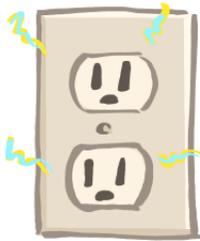


Figure 1a. Literal Target Picture

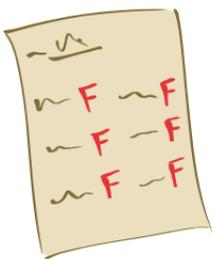


Figure 1b. Metaphor Target Picture

Literal versus Metaphor Biases in Materials

To determine whether the critical words were biased toward either one or the other of the relevant interpretations, a norming study was conducted with 48 new participants drawn from the same pool as the main experiment. These participants were given the initial sentence frame, ending before the disambiguating phrase, (e.g., “The actress was

huge because of the--”), and asked to finish the sentence with the first sensible completion that came to mind. Responses were subsequently coded as involving either a literal or a metaphorical reading of the adjective. A median split of items divided the ambiguous adjectives into those deemed Literal Biased (77% of responses presumed a literal interpretation; SD=26%) versus those deemed Metaphor Biased (15% of responses presumed a literal interpretation; SD=15%). The preferred reading may depend on the relative frequencies of literal versus metaphorical meanings of the adjectives in the language, and/or linguistic and extra-linguistic knowledge regarding the combination of the subject noun phrase and the critical word.

Procedure

Participants were instructed to listen carefully to the utterances, as there would be a memory test at the end of the study. Other than listening to and looking at the stimuli, and the memory test at the end, there was no task.

Participants’ eye movements were monitored using an Eyelink table-mounted eye tracker, sampling at 500 Hz. Images were presented on a 19-inch LCD flat screen monitor at a resolution of 1280 x 1024 pixels while spoken sentences were presented through a pair of headphones worn by the participants. For each trial, the images appeared 1,000ms before sentence onset. Viewing was binocular, but only one eye was tracked. Although head position was not physically restricted, participants were encouraged to minimize movements of the head.

Analysis

Analysis involved repeated measures ANOVAs on empirical log odds of fixations in the one-second time period prior to the onset of the disambiguating word (plus a 200-ms positive offset to allow time to plan and execute an eye movement; see, e.g., Salverda, Kleinschmidt, and Tanenhaus, 2014), that is, from -800 ms to +200 ms. Factors were Picture Category (target, competitor, average of unrelated pictures) x Preferred Reading (Literal, Metaphor) x Sentence Completion (completed as literal, completed as metaphorical). For Picture Category, we use “target” to refer to the picture that is favored by the eventual sentence completion, while “competitor” is the picture related to the alternate interpretation. Thus, a literal target will have a metaphorical competitor, and vice versa. For simplicity, unambiguous control trials are not discussed in the following analyses.

Results and Discussion

Before the Disambiguating Nouns

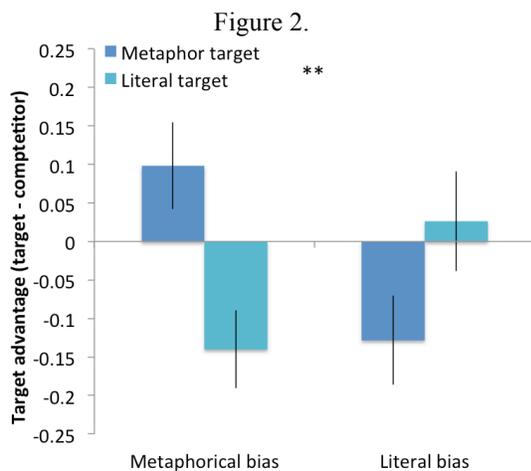
Initial analysis of the first interval revealed a significant effect of Picture Category ($F(2,74) = 9.04, p = .0003$), attributable to overall greater looks to targets ($t(37) = 3.45, p = .001$) and competitor pictures ($t(37) = 4.58, p < .0001$)

than to unrelated pictures. However, looks to targets and competitors did not differ from each other, as one might expect if listeners entertain both the literal and the metaphorical interpretations of the ambiguous adjectives. However, the main effect of Picture Category was qualified by a significant 3-way interaction with Preferred Reading and Sentence Completion ($F(2,74) = 7.54, p = .001$). No other effects or interactions reached significance.

To assess the nature of the three-way interaction, the effects of Preferred Reading and Sentence Completion were analyzed for each Picture Category. There were no significant effects for unrelated pictures. By contrast, competitor looks showed a Preferred Reading x Sentence Completion effect ($F(1,37) = 4.31, p = .04$), as did target looks ($F(1,37) = 10.72, p = .002$).

For competitor pictures, the interaction between Preferred Reading and Sentence Completion resulted because there were more looks to the metaphorical competitor on literal trials when the adjectives' preferred reading was metaphorical than to the literal competitor on metaphorical trials ($t(37) = 3.09, p = .004$), with a non-significant tendency in the opposite direction for trials with a literal preferred reading. The story was much the same for targets: when the preferred reading was metaphorical, there were more looks to metaphorical targets than to literal targets ($t(37) = 2.70, p = .01$), but when the preferred reading was literal, there were more looks to literal targets than to metaphorical targets ($t(37) = 2.55, p = .015$).

The interaction can be seen more readily in Figure 2, which shows the difference between the proportion of looks to the target and looks to the competitor for each of the four kinds of ambiguous items during the 1s interval immediately preceding the onset of the disambiguating information. Positive values on the difference scores represent more looks to the target, while negative values represent more looks to the competitor. The reliable interaction between Preferred Reading and Sentence Completion results because of a strong preference for the metaphor pictures in the metaphor biased sentences, and a somewhat weaker preference for the literal pictures in the literal biased sentences.



These analyses of the interaction term suggest a slightly different story than that of the simple, balanced competition posited on a parallel activation account. Specifically, there are strong effects of preferred reading, such that listeners tend to fixate the literal target for literal preferred readings, and the metaphorical target for metaphorical preferred readings. They then appear to switch readings if the disambiguating noun indicates the dispreferred alternative is correct.

After the Disambiguating Nouns

Given that listeners did not make overt responses, it was important to verify that they did eventually converge on the intended interpretation of each sentence. Accordingly, we conducted another ANOVA on empirical log odds of fixations to pictures in the 1-second interval *following* the onset of the disambiguating word. The ANOVA had the same factors as above. Here, Picture was again significant ($F(2,74) = 136.6, p < .0001$). Competitor looks again outpaced unrelated looks ($t(37) = 2.17, p = .03$), and target looks outpaced both overall (unrelated: $t(37) = 12.37, p < .0001$; competitors; $t(37) = 11.96, p < .0001$). Greater target than competitor looks implies that listeners were robustly fixating intended target pictures after hearing the disambiguating phrase.

However, as in the interval preceding the onset of the disambiguating information, the main effect of Picture Category was qualified by a significant three-way interaction with Preferred Reading and Sentence Completion ($F(2,74) = 7.54, p = .001$). Thus we cannot conclude that targets necessarily outpaced competitors in all four cells of Preferred Reading x Sentence Completion.

To assess this, we compared target looks and competitor looks in each cell, finding substantially greater target looks in all of them (preferred literal, literal completion: $t(37) = 8.21, p < .0001$; preferred literal, metaphoric completion: $t(37) = 5.31, p < .0001$; preferred metaphoric, literal completion: $t(37) = 8.04, p < .0001$; preferred metaphoric, metaphoric completion: $t(37) = 9.78, p < .0001$). This indicates that, in all conditions, listeners resolved the literal or metaphorical interpretations as intended.

Time Course of Adjective Interpretation

Figure 3 shows the proportion of looks to each of the picture types as participants listened to sentences with a preferred literal reading of the adjective. Time is aligned to the onset of the disambiguating word, so that the graph spans from 2 seconds before that point until 2 seconds after. The top panel shows that for the literal biased items (“The young woman was feeling sore”), looks to the literal target pictures (e.g., woman climbing on a cliff) exceed looks to the unrelated items before the onset of the disambiguating information “climb” (upper left panel). Interestingly, looks to the unbiased metaphorical competitor pictures (red line in the upper left panel) also exceed looks to the unrelated items, suggesting brief competition between the two interpretations. After the disambiguating noun, looks to the

target rapidly increase, and looks to the competitor drop to baseline.

The lower panel of Figure 3 shows looking behavior in literal biased items that ended with a metaphorical interpretation of the adjective. Here looks to the literal competitor pictures (e.g., woman climbing on a cliff) exceeded looks to the unrelated items before the onset of the disambiguating information (“argument”), and dropped off shortly thereafter. Likewise, looks to the unbiased metaphorical targets (e.g., a couple arguing) did not exceed looks to the unrelated targets until after the onset of the disambiguating noun (“argument”).

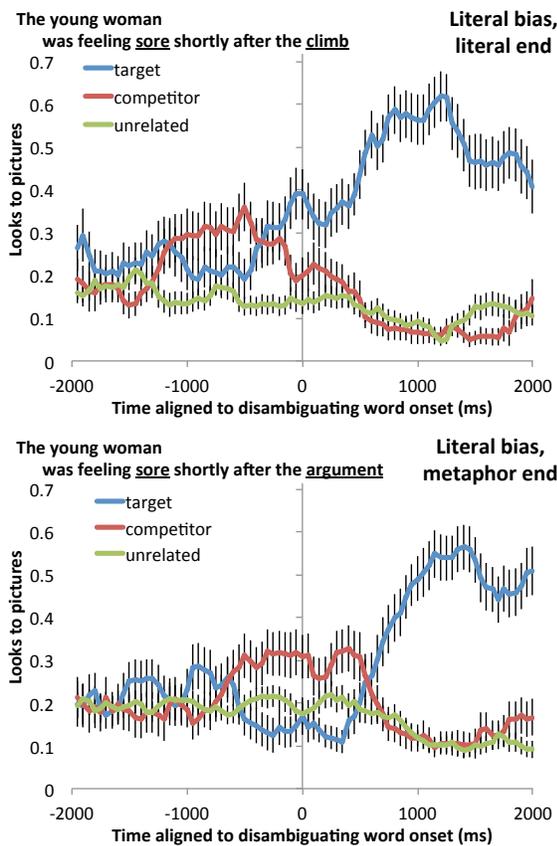


Figure 3.

Figure 4 shows looks to pictures as participants listened to metaphor biased items, such as “The lawyer was stunned,” with dispreferred targets in the top panel, and preferred targets in the bottom panel. When the sentence ended with the dispreferred literal interpretation (upper panel), participants initially look at the metaphor competitor (a judge’s gavel), and looks to the target picture (a bald man wielding a stick) began to increase only *after* the onset of the disambiguating information (“beating”). When the sentence ended with the preferred metaphorical interpretation, looks to the metaphorical target pictures (e.g., a judge’s gavel) exceed looks to unrelated items well before the onset of the disambiguating information (“verdict” at

time 0), and increased rapidly thereafter. In the case of the metaphor biased items, participants initially adopted the preferred interpretation of the ambiguous adjective, and abandoned it only after they encountered inconsistent information.

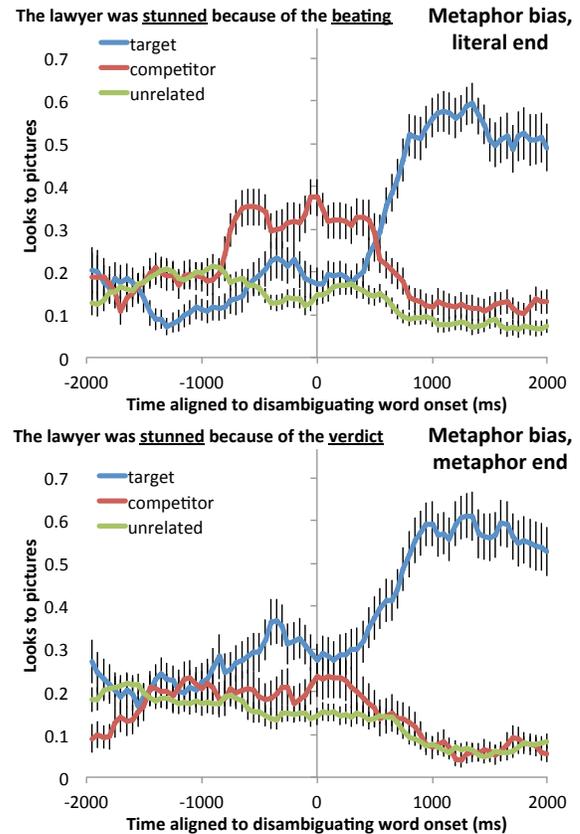


Figure 4.

Overall, these data provide very little support for parallel activation of literal and metaphorical interpretations of these adjectives (with the exception of the literal bias, literal ending items). Rather, participants initially adopt the contextually biased interpretation, and shift to the dispreferred interpretation only after the advent of disambiguating information.

Models of Metaphor Comprehension

Below we briefly consider the degree of fit between data from the present study and prominent models of metaphor comprehension.

Literal First Models The present study provides little support for Gricean models of metaphor comprehension in which listeners initially retrieve the literal interpretation of words, computing metaphoric interpretations only when literal ones are found to be anomalous. While our participants did show an early preference for the literal interpretation of adjectives appearing in our literal biased sentence contexts, their looking behavior during the metaphor biased items suggested the metaphoric

interpretations were initially adopted, with no evidence that participants ever activated the literal interpretation of adjectives. Results of the present study are thus in keeping with a large body of literature undermining these models (see Gibbs, 1994 for a review).

Parallel Activation Models of metaphor comprehension inspired by research in cognitive linguistics suggest that literal meanings play a functional role in the interpretation of metaphors, and thus predict that listeners should activate both readings – at least briefly. As noted above, the present study suggests participants commit to the preferred interpretation relatively early, and show little evidence for the activation of the dispreferred alternative until after the onset of disambiguating information.

Graded Saliency Given that we found evidence that listeners commit to a single particular interpretation, findings reported here are somewhat compatible with versions of the graded saliency model (GSM) that incorporate a separate non-modular context-sensitive mechanism (Giora, 2003). One wrinkle is that the GSM is committed to the proposal that context-driven facilitation of less salient meanings (as in the metaphors) cannot be done at a cost to the access of salient (viz. literal) meanings, so that our failure to find evidence for the activation of literal meanings in metaphor biased contexts is difficult to accommodate on this model.

Constraint Satisfaction Models Results of the present study can be accommodated quite well by models suggesting the computation of conventionalized metaphoric meanings is subject to the same sorts of contextual factors that impact the interpretation of literal meanings (e.g., Katz & Ferretti, 2001; MacDonald, Pearlmutter, & Seidenberg, 1994). In constraint satisfaction models, different sources of information (e.g., syntactic, lexical, conceptual) compete for activation in parallel over time, and constraints interact to provide probabilistic evidence for various potential interpretations. Such models depend crucially on the strength of the constraints, and consequently can account for data that involve competition between multiple meanings as well as data, like those reported here, in which some potential interpretations are not activated at all. Observed evidence for the early activation of preferred interpretations is especially compatible with ‘maximally incremental’ models that employ predictive coding, continuously combining linguistic and non-linguistic information in the dynamic computation of meaning (e.g., Altmann & Mirkovic, 2009).

Conclusion In summary, we show that listeners look at images compatible with the contextually biased interpretation of ambiguous adjectives, irrespective of whether the preferred reading is literal or metaphorical, contra metaphor comprehension models that posit the obligatory activation of literal meanings.

Acknowledgments

Thanks to Teresa Patty, William Ni, and Rachel Bristol for help with data collection.

References

- Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247-264.
- Altmann, G., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science*, 33(4), 583-609.
- Coulson, S., & Matlock, T. (2001). Metaphor and the space structuring model. *Metaphor and Symbol*, 16(3-4), 295-316.
- Coulson, S., & Oakley, T. (2005). Blending and coded meaning: Literal and figurative meaning in cognitive semantics. *Journal of Pragmatics*, 37(10), 1510-1536.
- Frisson & Pickering, 1999 Frisson, S., & Pickering, M. J. (1999). The processing of metonymy: evidence from eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(6), 1366.
- Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding*. Cambridge University Press.
- Gibbs, R. W., Bogdanovich, J. M., Sykes, J. R., & Barr, D. J. (1997). Metaphor in idiom comprehension. *Journal of memory and language*, 37(2), 141-154.
- Giora, R. (2003). *On our mind: Saliency, context and figurative language*. Oxford, England: Oxford University Press.
- Glucksberg, S., Gildea, P., & Bookin, H. B. (1982). On understanding nonliteral speech: Can people ignore metaphors?. *Journal of verbal learning and verbal behavior*, 21(1), 85-98.
- Grice, H. P. (1975). Logic and Conversation in P. Cole and J. Morgan (eds.) *Syntax and Semantics Volume 3: Speech Acts*.
- Katz, A. N., & Ferretti, T. R. (2001). Moment-by-moment reading of proverbs in literal and nonliteral contexts. *Metaphor and Symbol*, 16(3-4), 193-221.
- Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to western thought*. Basic books.
- MacDonald Pearlmutter & Seidenberg (1994) MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological review*, 101(4), 676.
- Richardson, D., & Matlock, T. (2007). The integration of figurative language and static depictions: An eye movement study of fictive motion. *Cognition*, 102(1), 129-138.
- Salverda, A. P., Kleinschmidt, D., & Tanenhaus, M. K. (2014). Immediate effects of anticipatory coarticulation in spoken-word recognition. *Journal of memory and language*, 71(1), 145-163.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268(5217), 1632-1634.