

BUCLD 36 Proceedings
To be published in 2012 by Cascadilla Press
Rights forms signed by all authors

Information Packaging for Causative Events: Implications for Crosslinguistic Production and Attention

Ann Bunger, John C. Trueswell, and Anna Papafragou

1. Introduction

When speakers are describing events, in particular events that involve complex interactions between multiple entities, they have to make choices about *what* information about a given event they want to convey and *how* to encode that information in language. In addition to their communicative intent, the way that speakers describe events is influenced by both perceptual-conceptual factors and language-specific lexical and syntactic factors that determine the way that event information can be packaged into linguistic units. In the current study, we investigate the way that speakers package information about complex causative events into event descriptions, with an eye to both crosslinguistic and developmental differences in causative event description.

In a causative event, an Agent performs an action on a Theme object which, in turn, causes some change of state or location in the Theme. Imagine, for example, the moment in a soccer game in which a player kicks the ball, thereby sending the ball across the field and into the goal. There is theoretical and experimental evidence that both children and adults conceive of causative events like this as being composed of two distinct subevents: a Means and a Result (e.g., Bunger, 2006; Cohen, Rundell, Spellman & Cashon, 1999; Jackendoff, 1990; Talmy 1985). The Means subevent corresponds to the action of the Agent on the Theme object: in this example, the player's kicking of the ball. The Result subevent corresponds to the change that the Theme undergoes: in this example, the movement of the ball to the goal. Moreover, humans interpret these interactions as subevents of a complex causative event rather than as a sequence of unrelated happenings because of a perceived causal link between the Means and Result. That is, we understand that the player's kicking of the ball directly resulted in the ball's change of location.

When describing a causative event, a speaker has to make choices about how much of the event, i.e., how many subevents, to talk about, and how to encode that information in an utterance. In addition, the particular language that

* Ann Bunger, University of Delaware, bunger@udel.edu; John C. Trueswell, University of Pennsylvania, trueswel@psych.upenn.edu; Anna Papafragou, University of Delaware, papafragou@psych.udel.edu. This research was supported in part by Grant #R01-HD055498 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development to A. Papafragou and J.C. Trueswell.

a speaker is using places constraints on the way that event information can be packaged in linguistic structures. These linguistic factors affect what information about an event is likely to get encoded as well as how that information is mapped to language (e.g., Talmy, 2003). In both English and Greek, for example, speakers have the option of encoding information about only one causative subevent in an utterance without providing information about the other. However, Greek places constraints that English does not on the way speakers can encode information about *both* causative subevents.

The sentences in (1) provide an example of how English (1a) and Greek (1b) speakers might describe only the Means of our example causative event. The verb “kick” (Greek “klotso”) describes the nature of the player’s action on the ball, but the sentence does not provide any information about what happened to the ball after it was kicked. Likewise, the sentences in (2) provide information only about the Result subevent. The verb “sent” (Greek “estile”) describes an Agent-driven transfer of location, and the endpoint of the ball’s trajectory is encoded in the prepositional phrase “into the goal” (Greek “sto terma”), but the sentence does not provide any information about how the ball arrived at this location or what the player did to get it there. Based on the sentences in (2), we do not know whether the player used a foot, a lacrosse stick, or the Force to move the ball.

(1) a. The player kicked the ball.

b. O pektis klotsise ti bala.
The player kicked the ball

(2) a. The player sent the ball into the goal.

b. O pektis estile ti bala sto terma.
The player sent the ball into the goal

English also offers a standard recipe for packaging both causative subevents into a single sentence: English speakers can combine a verb like “kick” that specifies a Means activity with a post-verbal prepositional phrase like “into the goal” that specifies a Result, giving rise to something like the sentence in (3a). However, because of constraints on the use of resultative phrases that describe a bounded event (Giannakidou & Merchant, 1999; Papafragou, Hulbert & Trueswell, 2008; Papafragou & Selimis, 2010), Greek does not allow speakers this option of efficiently packaging Means and Result information in the same clause. The Greek equivalent of the sentence in (3a), shown in (3b), is ungrammatical, and instead, a Greek speaker would describe both causative events by stringing together multiple sentences, each of which encodes a single subevent, as in (3c).

(3) a. The player kicked the ball into the goal.

b. * O pektis klotsise ti bala sto terma.
The player kicked the ball into the goal

c. O pektis klotsise ti bala ke afti pige sto terma.
The player kicked the ball and it went into the goal

Note that although the Greek method of encoding both subevents is just as informative as the English method, it is not as efficient in terms of linguistic resources, and it also is not a standard packaging strategy in the way that the English option is. Because Greek speakers do not have a formula to follow when describing a causative event, it is likely that they experience more competition among a wider range of information packaging alternatives than English speakers do.

Given these crosslinguistic differences in the way that information about causative events can be packaged in linguistic structures, our goal was to investigate what effect these differences have on the way that children and adults describe causative events. First, we want to ask how speakers of English and Greek package information for causative events. In particular, we want to find out how speakers of Greek resolve this constraint on packaging both causative subevents into a single clause. In addition, we want to ask how the description of causative events changes with development. In general, children are more likely than adults to be underinformative when describing complex events. Indeed, Bungler, Trueswell and Papafragou (2012) demonstrated in an eyetracking study that children tend to omit details from their descriptions of complex events that adults include even when children have paid as much attention to those details as the adults have. Given this tendency toward underinformativeness, and coupling it with the fact that children do not process language as efficiently as adults, we want to find out how children decide what to say about causative events.

In this study, we assess both the descriptions that speakers provide for causative events and the attention that they pay to various components of those events. A speaker's attention relates to their language production in both trivial and non-trivial ways. Trivially, speakers are more likely to talk about things that they have paid attention to, rather than things that they haven't paid attention to. Less trivially, speakers are likely to map event components to language in the same order that they attended to them (e.g., Gleitman, January, Nappa & Trueswell, 2007; Griffin & Bock, 2000). Here, we use eyegaze during language production as a window into the process that speakers go through while selecting which event components to encode in language. Trivially, we expect speakers to pay more attention to the subevents that they mention, but what is equally interesting is the way that speakers direct their attention when they omit information from their event descriptions. On the one hand, speakers might fail to look at subevents they do not talk about, suggesting that they have not considered those details at all. On the other hand, if speakers are dealing with competition between different packaging options, we might actually expect them

to devote a significant amount of attention to event components that they do not end up talking about.

2. Methods

2.1 Participants

Data were collected from native monolingual speakers of English and Greek in three age groups: 3-year-olds (English: $n=17$, mean age 3;5, range 3;0–3;12; Greek: $n=20$, mean age 3;7, range 3;2–3;12), 4-year-olds (English: $n=19$, mean age 4;6, range 4;1–4;12; Greek: $n=20$, mean age 4;7, range 4;0–5;1), and adults (English: $n=19$; Greek: $n=18$). Parents of English-speaking children were recruited through preschools in Newark, DE and Philadelphia, PA; English-speaking adults were students at the University of Pennsylvania and received course credit for participation. Parents of Greek-speaking children were recruited through preschools in and around Ioannina, Greece; Greek-speaking adults were students at the University of Athens and received monetary compensation for participation.

2.2 Materials

The stimuli consisted of still clip-art images that depicted the midpoint of various events. Target items depicted Agent-driven causative events that resulted in a change of location for the Theme object. In each target event, an animate Agent used some Instrument (a tool or body part) to interact with an inanimate Theme object (the Means subevent) in a way that caused the Theme to move toward a visible Goal (the Result subevent). A sample target item is provided in Figure 1: in this event, the boy has used his fist to punch the soccer ball (Means), thereby sending the soccer ball (in)to the basket (Result). Clipart images were constructed such that the tool or body part used as an Instrument was separated spatially from the face of the Agent, allowing looks to these two regions to be distinguished in the analysis of eyegaze data. Filler pictures depicted animate Agents involved in events in which they did not cause a Theme object to move from one place to another (e.g., a snail and a rhinoceros playing trumpets). The full set of 8 target items and 15 fillers were presented to adult participants; children were presented with a subset of these items, 6 target pictures and 10 fillers.

2.3 Procedure

Each participant was run individually at his/her university campus or preschool. The consent process and experiment instructions were carried out in each participant's native language.

Stimuli were presented to all participants on the same Tobii 1750 remote eyetracking system. The data sampling rate was a consistent 50 Hz, and spatial resolution of the tracker was approximately 0.5–1 degrees of visual angle,

including corrections for small head movements. Participants were seated approximately 60 cm from the Tobii screen, and experimenters adjusted the angle of the screen as necessary to obtain robust views of both eyes, centered in the tracker's field of view. All participants were calibrated using the ClearView software's default 5-point calibration scheme.



Figure 1—Sample target event

During the experiment, participants viewed a sequence of stimulus items (23 items for adults, 16 items for children) presented in one of two fixed semi-random orders. While viewing these stimuli, participants performed one of two randomly-assigned between-subjects tasks. Half of the participants in each age group were asked to study each of the pictures carefully in preparation for a memory game¹ (Nonlinguistic condition), and the other half were also asked to provide a description of each picture when it appeared on the screen (Linguistic condition). Each item was displayed for a limited amount of time, and a beep sounded as each item appeared as a cue to begin either examining (Nonlinguistic condition) or describing (Linguistic condition) the picture. Responses were provided verbally, and sessions in which participants completed the Linguistic condition were audio recorded to facilitate accurate collection of event descriptions.

2.4 Data analysis

¹ Stimuli and procedure for the Memory task are not described in this paper because they are not relevant to the questions under investigation. The Memory task was presented to each participant after he or she had completed one of the Linguistic or Nonlinguistic tasks described in the text, and thus it did not interfere with collection of the data reported here.

Event descriptions collected from participants in the Linguistic condition were transcribed and hand-coded by native speakers of the language under consideration. Descriptions of target items were assessed for mention of the two subevents of each causative target. Words and phrases that referred either to the Agent's activity or to the nature of the Agent's contact with the Theme object were coded as mention of the Means subevent. Words and phrases that identified either what happened to the Theme object as a result of that activity or some Goal-oriented intention on the part of the Agent were coded as mention of the Result subevent. For example, for the target event depicted in Fig. 1, participants used verbs like "punch," "throw" (Greek "rhino"), and "hit" ("htipo") to describe the Means subevent, and verbs like "put" ("bazo") and prepositional phrases like "into the basket" ("sto kalathi") to describe the Result subevent.

Eye movement data were analyzed to assess the effects of condition, age, and language background on mention of causative subevents for each target item. Data samples (50 per second) were timelocked to the onset of each stimulus item. Trackloss was determined separately for each eye by Tobii's ClearView software. Gaze coordinates were taken from eyes with no trackloss (or from an average of both eyes, if neither had trackloss). Target trials with global trackloss of >33% were excluded from the analysis ($n=27$ for 3yos, $n=25$ for 4yos, $n=9$ for adults). Trials were excluded evenly across target items for each age group. In addition to the participants described in Section 2.1, one participant from each language and age group who had more than four excluded target trials was excluded from the analysis.

Regions of interest (ROIs) that corresponded to the Means and Result subevents of each target picture were defined on the basis of independent norming by adult raters from each language group ($n=12$ English speakers, $n=6$ Greek speakers). Raters were asked to draw rectangles around the part of each causative target picture that showed "the Result of the event" and the part that showed "the Central Action or Means." ROIs were based on the most common region assignments across items: looks to the Instrument used by the Agent to interact with the Theme object were interpreted as looks to the Means subevent, and looks to the Goal object were interpreted as looks to the Result subevent. For the event depicted in Fig. 1, then, looks to the boy's fist were interpreted as looks to the Means subevent and looks to the basket were interpreted as looks to the Result subevent. Note that the Theme object (the ball, in Fig. 1) was not included in either ROI. Although Themes are conceptually critical to both causative subevents, our raters were more likely to omit them from each region than to include them. Means and Result ROIs never overlapped in spatial coordinates.

3. Results

3.1 Event descriptions

The descriptions provided for target events confirmed our predictions about the effects of crosslinguistic and developmental challenges to information packaging, namely, that speakers would deal with effortful information packaging situations by omitting details about the events they were describing. In general, Greek speakers mentioned fewer causative subevents than English speakers, and in both language groups, children mentioned fewer causative subevents than adults.

Table 1 provides information about the number of causative subevents mentioned in event descriptions. Across language and age groups, speakers were significantly less likely to mention both subevents of a given event than they were to omit one or both. When participants did mention both subevents, they always mentioned the Means first. English speakers tended to encode the Means subevent in a verb and the Result in a post-verbal prepositional phrase, as in the sentence in (4a); Greek speakers tended to provide a verb that encoded the Means subevent in the first of two phrases, followed by phrase including a verb that encoded the Result subevent, as in the sentence in (4b).

- (4) a. A person just punched a soccer ball into a basket. adult
 b. Enas antras htipai mia bala gia na ti vali sto kalathi. adult
 A man is-hitting a ball so as to put-it into the basket

In about a third of the event descriptions, speakers mentioned only one causative subevent. The sentences in (5), which encode just the Means subevent, and (6), which encode just the Result subevent, were provided as descriptions of the target event in Fig. 1.

- (5) a. He's punching a soccer ball. 4-year-old
 b. Anthropos pu erikse ti bala. 4-year-old
 man who threw the ball
- (6) a. He's trying to get the soccer ball into the basket. 3-year-old
 b. Prospathi na vali afti ti bala mesa edo. 4-year-old
 trying to put this DET ball in here

In a surprisingly high number of event descriptions, speakers did not mention either of the causative subevents. Instead, for these trials, they described target events at a superordinate level, as in the sentences in (7), a strategy that might provide a kind of "escape hatch" from competition between packaging alternatives.

- (7) a. He's playing ball. 3-year-old

b. Enas babas pu pezi bala.
 a dad who plays ball

3-year-old

Moreover, event descriptions demonstrate effects of both language background and age on the number of subevents that participants mentioned (Table 1). Overall, English speakers were more likely than Greek speakers to mention both subevents. Within both language groups, however, adults were more likely than children to mention both subevents. English speakers were also more likely than Greek speakers to mention one subevent, but this effect of language background was driven by the English-speaking children who were most likely to mention only one subevent. Finally, Greek speakers were more likely than English speakers to mention neither of the causative subevents, and in both language groups, children were more likely than adults to mention neither subevent.

Table 1—Proportion of mention of Means and Result subevents in descriptions of causative event targets

	Both subevents	One subevent	Neither subevent
All participants	0.26 (\pm 0.04)*	0.35 (\pm 0.03)	0.39 (\pm 0.04)
English			
Adults	0.66 (\pm 0.06)	0.20 (\pm 0.04)	0.14 (\pm 0.03)
4-year-olds	0.20 (\pm 0.04)	0.52 (\pm 0.06)	0.28 (\pm 0.07)
3-year-olds	0.07 (\pm 0.03)	0.56 (\pm 0.11)	0.37 (\pm 0.11)
Greek			
Adults	0.39 (\pm 0.10)	0.29 (\pm 0.07)	0.32 (\pm 0.09)
4-year-olds	0.03 (\pm 0.03)	0.40 (\pm 0.09)	0.57 (\pm 0.08)
3-year-olds	0.06 (\pm 0.04)	0.20 (\pm 0.07)	0.74 (\pm 0.09)

* $p < 0.05$ vs. One subevent and Neither subevent.

We tested the reliability of these observations using multilevel mixed logit modeling with crossed random intercepts for Subjects and Items (after Baayen, Davidson & Bates, 2008). Categorical values at the trial-level for number of

subevents mentioned in event descriptions (0, 1, 2) were modeled using Language (English, Greek) and Age (Adult, 4-year-old, 3-year-old) as (between-subjects) first-level fixed factors. The best-fitting model revealed main effects of both Language and Age, but no interaction between the two ($p < 0.001$ vs. an empty model with no fixed effects). The fixed effects reported in Table 2 reflect the fact that the effect of Age was driven by differences between adults and children in each language group, but not by differences between 3-year-olds and 4-year-olds.

Table 2—Fixed effects from best-fitting multilevel linear models of subevent mention for each subevent packaging strategy

Effect	Estimate	S.E.	t-value
Intercept	0.59	0.17	3.36
Language (English vs. Greek)	-0.44	0.10	-4.52
Age (3yo vs. 4yo)	0.19	0.12	1.58
Age (3yo vs. adult)	0.93	0.12	7.67

Formula in R: $\text{DepVar} \sim \text{Language} + \text{Age} + (1 \mid \text{Subject}) + (1 \mid \text{Event})$

3.2 Eye movements

Examination of eye movements by participants in the Linguistic condition provides insight into the processes that speakers were going through as they prepared event descriptions. A comparison between eye movements in the Nonlinguistic condition (Fig. 2) and the Linguistic condition (Fig. 3) support the established finding that the process of language production shifts a speaker's attention away from their baseline preferences. In this study, these shifts in attention during language planning play out differently for different groups of speakers. Paralleling the trends we saw in event description, we found separate effects of both language background and age on eyegaze during event description. Although we found more competition for attention between regions of our target images that depicted each of the causative subevents for Greek speakers than for English speakers, in general, adults allocated attention to both causative subevents even when they ended up mentioning only one. We did not observe this same pattern of eyegaze in children, however: children who mentioned only one causative subevent attended less to the one they failed to mention, suggesting that they did not actually consider it for inclusion in their utterance.

Figure 2 shows eyegaze patterns for participants in the Nonlinguistic condition, which provide information about baseline preferences for Means and Result regions of the target images. Attention allocation was equivalent across all age and language groups in this condition: participants showed no early preference for either causative subevent, but as they continued to view the event, their attention shifted toward the endpoint.

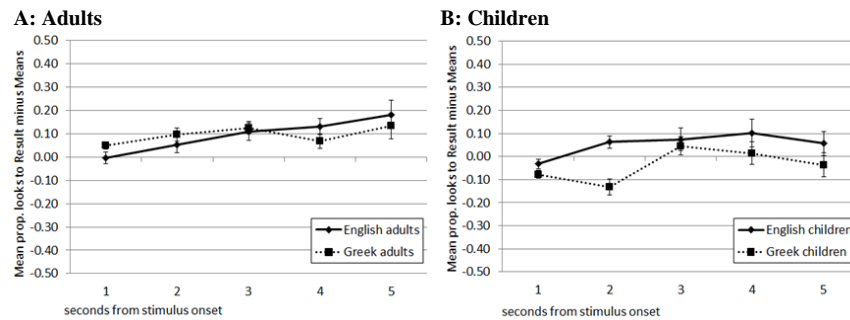
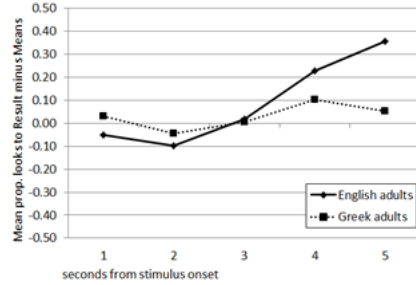


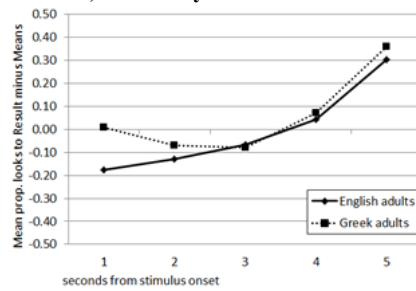
Figure 2—Mean proportion of looks to causative subevent regions by adults (A) and children (B) in the Nonlinguistic condition. Values represent difference scores (Result looking minus Means looking): values above zero correspond to looks at the Result subevent, and values below zero correspond to looks at the Means subevent. Gaze preferences are averaged across 1-sec units of the viewing period.

Figure 3 shows eyegaze patterns for participants in the Linguistic condition, split by the number of causative subevents mentioned in event descriptions. For adults who mentioned both causative subevents (Fig. 3A), there are clear differences in attention allocation across language groups. English-speaking adults who mentioned both causative subevents shifted their attention between the two subevents during language production, showing an early preference for the Means region in comparison to participants in the Nonlinguistic condition and then shifting their gaze to the Result region. This order of attention to the two regions directly corresponds to the order in which speakers mentioned the subevents in their event descriptions: Means first, and then Result. Greek adults who mentioned both subevents also exhibited an early preference for the Means region. Unlike the English speakers, however, within the timeframe presented in Fig. 3A, Greek speakers did not shift their attention to the Result region. This suggests that for Greek adults, these two subevents were still in competition for their attention.

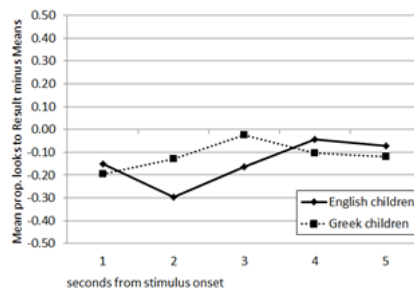
A: Adults, Both subevents



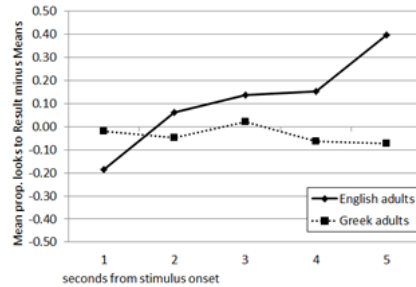
B: Adults, Means only



C: Children, Means only



D: Adults, Result only



E: Children, Result only

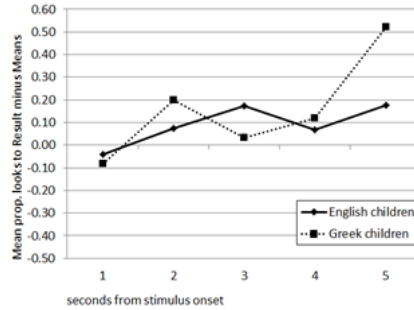


Figure 3— Mean proportion of looks to causative subevent regions in the Linguistic condition. Data are presented separately for adults who mentioned both causative subevents (A), adults (B) and children (C) who mentioned only Means subevents, and adults (D) and children (E) who mentioned only Result subevents. Values represent difference scores (Result looking minus Means looking), as in Fig. 2. Gaze preferences are averaged across 1-sec units of the viewing period. Data is collapsed across age groups for children because there were no differences between age groups in patterns of event description. Data for children who produced both subevents was quite sparse, and is not included.

Adult speakers of both languages who mentioned only the Means subevent (Fig. 3B) exhibited a preference for the Means region early in the trial and then shifted their gaze to the Result region. On the contrary, children from both language groups who mentioned only the Means subevent (Fig. 3C) showed a preference for the Means region throughout this viewing period. For adults who mentioned only the Result subevent (Fig. 3D), we again observed differences in eyegaze patterns between the two language groups: English speaking adults shifted their attention early to the Result region, but Greek speakers again demonstrated evidence of competition between the two subevents for their attention. Finally, just as children who mentioned only the Means subevent looked more at the Means region, children in both language groups who mention only the Result (Fig. 3E) looked more at the Result region.

4. Conclusions

When faced with situations that posed challenges for event description, speakers in this study tended to sacrifice informativeness for the sake of efficient information packaging. Specifically, they omitted details about causative events when packaging them in language required more effort. We saw both crosslinguistic and developmental effects on information packaging. Across age groups, Greek speakers were more likely than English speakers to omit subevents from their descriptions of causative events, an observation that can be connected to the fact that Greek speakers experience more competition between information packaging alternatives than English speakers do. Across language groups, the ability to efficiently package information about complex events improved with development, such that adult speakers of both languages tended to mention more causative subevents than children did.

Finally, the eye movement trends observed in this study confirmed that the process by which speakers decided what to talk about changed with development. Specifically, the results demonstrate that children and adults were omitting details about our events for different reasons. Although children and adults examined our events in the same way when they were not describing them, when they were describing them, children's attention was more restricted, and younger speakers ended up talking about the parts of an event that they had spent the most time looking at. Adults, on the other hand, seem to have weighed all the subparts of an event against language-specific constraints on information packaging to make conscious choices about how to package information about that event in their description of it. The fact that eyegaze in the Nonlinguistic condition was equivalent across age and language groups confirms that differences between groups of speakers in the Linguistic condition reflect the way that each group dealt with the added processing costs associated with language planning, rather than being characteristic of fundamental differences in event inspection or conceptual event encoding across the groups.

References

- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412.
- Bunger, Ann. (2006). How we learn to talk about events: Linguistic and conceptual constraints on verb learning. Unpublished Ph.D. dissertation, Northwestern University.
- Bunger, Ann, Trueswell, John C., & Papafragou, Anna. (2012). The relation between event apprehension and utterance formulation in children: Evidence from linguistic omissions. *Cognition*, 122, 135–149.
- Cohen, Leslie B., Rundell, Leslie J., Spellman, Barbara A., & Cashon, Cara H. (1999). Infants' perception of causal chains. *Psychological Science*, 10, 412–418.
- Giannakidou, Anastasia, & Merchant, Jason. (1999). Why Giannis can't scrub his plate clean: On the absence of resultative secondary predication in Greek. In A. Mozor (ed.), *Greek Linguistics '97: Proceedings of the 3rd International Conference on Greek Linguistics* (pp. 93–103). Athens: Ellinika Grammata.
- Gleitman, Lila R., January, David, Nappa, Rebecca, & Trueswell John C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, 57, 544–569.
- Griffin, Zenzi. M., & Bock, Kathryn. (2000). What the eyes say about speaking. *Psychological Science*, 11, 274–279.
- Jackendoff, Ray. (1990). *Semantic Structures*. Cambridge, MA: MIT Press.
- Papafragou, Anna, Hulbert, Justin, & Trueswell, John C. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108, 155–184.
- Papafragou, Anna, & Selimis, Stathis. (2010). Lexical and structural biases in the acquisition of motion verbs. *Language Learning and Development*, 6, 87–115.
- Talmy, Leonard. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Grammatical categories and the lexicon* (Vol. III, pp. 57–149). Cambridge: Cambridge University Press.
- Talmy, Leonard. (2003). *Toward a cognitive semantics: Typology and process in concept structuring*. Cambridge, MA: MIT Press.