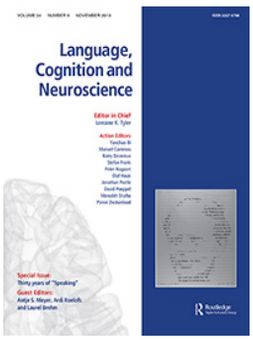

Language, Cognition and Neuroscience

ISSN: 2327-3798 (Print) 2327-3801 (Online) Journal homepage: https://www.tandfonline.com/loi/plcp21

The role of conceptualization during language production: evidence from event encoding

Anna Papafragou & Myrto Grigoroglou

To cite this article: Anna Papafragou & Myrto Grigoroglou (2019) The role of conceptualization during language production: evidence from event encoding, Language, Cognition and Neuroscience, 34:9, 1117-1128, DOI: 10.1080/23273798.2019.1589540

To link to this article: https://doi.org/10.1080/23273798.2019.1589540

LANGUAGE, COGNITION AND NEUROSCIENCE
2019, VOL. 34, NO. 9, 1117–1128
https://doi.org/10.1080/23273798.2019.1589540

Routledge
Taylor & Francis Group

REVIEW ARTICLE

The role of conceptualization during language production: evidence from event encoding

Anna Papafragou[a] and Myrto Grigoroglou[b]*

[a]Department of Psychological and Brain Sciences, University of Delaware, Newark, DE, USA; [b]Department of Linguistics and Cognitive Science, University of Delaware, Newark, DE, USA

ABSTRACT

As Pim Levelt proposed in *Speaking*, language production begins with the preverbal, conceptual apprehension of an event or state of affairs that the speaker intends to talk about. Despite the obvious importance of this process, relatively few studies to date have explored how conceptual representations are formed prior to speaking. Here we present a programme of research that tackles this question, focusing on the domain of events. Three key findings emerge. First, conceptual event structure shows important homologies with language. Second, given that event encoding differs across languages, the assembly of event representations prior to speaking varies cross-linguistically. Finally, conceptualising an event for purposes of communication depends not only on conceptual and linguistic factors but also on the pragmatic assessment of the needs and knowledge of the speaker's conversational partner. We sketch implications of this integrated approach to event conceptualisation for future research on how thought is transformed into language.

ARTICLE HISTORY
Received 1 October 2018
Accepted 25 February 2019

KEYWORDS
Conceptualization; events; production; message; motion

Introduction

Every act of speaking begins with an act of thinking. Building on this intuitive idea, in *Speaking*, Pim Levelt proposed that the first step in planning an utterance is conceptualisation – the preverbal apprehension of the broad content at the core of the speaker's communicative intention. Conceptualization forms the input to later processes of linguistic formulation and articulation and eventually results in a linguistic string (Levelt, 1989; see also Bock & Levelt, 1994; Bock, Irwin, & Davidson, 2004; cf. Lashley, 1951; Paul, 1970; Wundt, 1970). On this model, the level of conceptualisation contains a highly structured package of information (also called *message*) that the speaker wants to convey. This package of information draws from basic categories that the human mind uses to make sense of the world of experience, such as persons, objects, events, actions, states, times, places, directions and manners. *Speaking* observed that, at the time, a fully developed theory of the structure of messages was lacking but set out to offer "a global review of the message features that are required at later stages of processing, and indicate why they are needed" (Levelt, 1989, p. 70). According to the features identified by this account, messages are structured in propositional form, since they are the vehicles of reference and predication. Furthermore, they have thematic structure, in that they capture information about the entities that participate in an event or state of affairs, certain characteristics of those entities, and the relations among them (e.g. Agent, Goal, Instrument). Messages additionally specify information about temporal structure (past, present, future), the perspective from which information is presented, and other meaning components. Together these features are necessary and sufficient for the next stage of processing corresponding to linguistic formulation: during this later stage, the form of the utterance is determined as speakers select lexical items, assemble them into syntactic constituents, and engage in phonological and articulatory encoding.

This pioneering model has fuelled much subsequent empirical research and theorising on language production and its many sub-processes, including specific attempts to spell out the workings of the message formulation (see Konopka & Brown-Schmidt, 2014, for a recent review). However, thinking has remained a challenging topic to study and relatively few studies have addressed the contents of preverbal representations. At present, despite important progress in the field, several gaps remain in our understanding of the earliest stages

CONTACT Anna Papafragou Papafragou@psych.udel.edu
*Present address: Department of Applied Psychology and Human Development University of Toronto, Toronto, ON, Canada.

of the processes that transform thoughts into language. First, we know little about how the human mind constructs conceptual representations, and proposals (including those in *Speaking*) about both the kinds of representational units available to conceptualisation and the temporal profile of the processes that mobilise these units remain understudied. For instance, recall that one of the central properties of preverbal messages is considered to be their thematic structure that connects entities to their roles in an event or state of affairs (Levelt, 1989; cf. also Dowty, 1991; Jackendoff, 1990; Levin & Rappaport-Hovav, 2005; Pinker, 1989, a.o.). Nevertheless, direct evidence for this basic claim has until recently been lacking.

Second, we do not know how conceptual representations are mapped onto language. This issue becomes more complex if one considers that, across natural languages, the lexical-structural representation of space, number, objects and events varies considerably. *Speaking* strongly espoused the possibility that the conceptual information selection for purposes of speaking might be language-specific as the conceptualisation mechanisms become attuned to what is required by each linguistic system (a process also known as "thinking for speaking"; Slobin, 1996). However, hypotheses about how cross-linguistic variation might impact the mechanics of language production have until recently remained virtually untested. Furthermore, cross-linguistic differences raise the question whether the underlying conceptual representations might vary in the minds of speakers of different languages *even beyond the needs of speech preparation*, as people may habitually pay attention to language-congruent distinctions over others in a variety of contexts (see Bowerman & Levinson, 2001; Gentner & Goldin-Meadow, 2003; Gleitman & Papafragou, 2016 for different perspectives).

A final issue is that, as the cover of *Speaking* itself suggests, speaking typically happens in the context of conversation, since speech is usually initiated by the conception of some communicative intention. *Speaking* recognised that the generation of preverbal messages is firmly embedded within a communicative episode, in which the speaker wants to achieve some purpose by saying something, and wants the listener to recognise that intention from what was said. The speaker then selects how much and what type of information to encode so as to fulfil that conversational goal, and keeps track of the listener's needs and perspective to ensure that the selected message makes a relevant and informative conversational contribution (cf. Clark, 1996; Clark & Marshall, 1981). *Speaking* also suggested that the pragmatic attunement of message generation interacts with language-specific factors, since some semantic components (such as tense in some languages) are obligatory and have to be included in a message, whether they have communicative value or not (Levelt, 1989). Later empirical work has only selectively taken up the rich pragmatic processes surrounding the generation of messages that *Speaking* outlined. For instance, most current studies on how speakers select information to be encoded focus on a relatively restricted set of circumstances and tasks (mainly involving the identification of a single referent among several objects in an array; Krauss & Glucksberg, 1977; Krauss & Weinheimer, 1964), and very few incorporate cross-linguistic perspectives.

Here we present results from an ongoing, collaborative programme of research that addresses these open issues about conceptual structure and its interface with language along themes introduced and discussed in the pages of *Speaking*. This line of work combines linguistic tools with eye tracking and other online methods in novel ways to study the (otherwise inscrutable) conceptualizations that form the impetus for a production episode, as well as their rapid mapping to linguistic representations as speakers plan their utterances over time, and their adaptations to pragmatic factors. We focus on dynamic events that are ideally suited to addressing these issues: events in language have complex, internally structured representations that have been assumed to interface with – and probably reflect – similarly structured conceptual representations (Jackendoff, 1990; Shipley & Zacks, 2008); the lexical and morphosyntactic encoding of events is characterised by intense cross-linguistic variation (Bowerman & Choi, 2001; Majid, Boster, & Bowerman, 2008); and events can be described from multiple perspectives and at various levels of granularity depending on what the addressee knows or needs (Brown & Dell, 1987). We therefore use the domain of events to pursue three interlocking themes at the interface of thought and language: (a) How are events conceptualised in non-linguistic cognition? Is there concrete evidence for a homology between conceptual and linguistic event representation? (b) How are event representations mapped onto language during production? Are cross-linguistic differences in the content and structure of event descriptions reflected in the preparation of event messages? (c) Do speakers produce messages about events in a way that reflects the knowledge and goals of their addressees? Furthermore, do these messages differ depending on language-specific constrains on event encoding?

Conceptualizing events non-linguistically

Studies of event cognition commonly posit that, when encoding information about the complex and

continuous activity occurring in the environment, the human mind creates representations of events that capture abstract spatial, temporal, and causal information about the world (Radvansky & Zacks, 2014; Shipley & Zacks, 2008). Relatedly, as mentioned already, events as preverbal messages have been assumed to have a predicate-argument structure, including thematic roles that capture "who did what to whom" relationships in conceptual structure (Levelt, 1989; cf. Dowty, 1991; Jackendoff, 1990; Levin & Rappaport-Hovav, 2005; Pinker, 1989). These conceptual roles (expressed by corresponding linguistic constituents in sentences) include Agents (*A diver is swimming*), Patients (*A boy is kicking a ball*), Goals (*A butterfly flew to a flower*), Instruments (*A man is putting a pizza into the oven with a peel*), and so on. Some studies have suggested that information about individual event components or relationships between event components that determine whether an event is coherent or not can be extracted rapidly by human viewers – but these studies as a whole have focused on a limited number of events and event roles (see Dobel, Gumnior, Bölte, & Zwitserlood, 2007; Griffin & Bock, 2000, respectively; cf. also Dobel, Glanemann, Kreysa, Zwitserlood, & Eisenbeiss, 2010; Webb, Knott, & MacAskill, 2010; Zwitserlood et al., 2018).

To test whether event cognition is structured in a way that maps onto language, we asked whether people could recognise event categories (e.g. "pushing") and extract event roles such as Agents and Patients from very short visual displays of events (Hafri, Papafragou, & Trueswell, 2013). Participants were presented with photographs of 2-participant events (e.g. a girl pushing a boy) in 37 and 73 ms displays followed by a mask. They were then asked to answer forced-choice questions about Agents (e.g. *Is the girl performing the action?*), Patients (e.g. *Is the boy being acted upon?*), the event type (e.g. *Did you see pushing?*), or the particulars of the entire event (e.g. *The girl pushed the boy*). Results showed that participants were able to both extract thematic roles and recognise events and event categories even from the shortest displays. Further experimentation showed that, in order to extract thematic roles, people relied on visual cues typically associated with specific event categories and roles (e.g. an outstretched arm for the Agent of a pushing event); in a version of the experiment, where visual cues typically associated with agenthood were displayed by the patient (e.g. the Patient of a pushing event had an outstretched arm), people were less accurate in extracting thematic roles in short displays (but were very accurate in 2 s displays). Later work demonstrated that event roles are extracted not only rapidly but also spontaneously (i.e. involuntarily), even when people are asked to attend to event details that are irrelevant to event structure (Hafri, Trueswell, & Strickland, 2018). Together these findings suggest that people organise the flux of visual input into conceptually meaningful units rapidly and spontaneously; furthermore, this organisation reveals homologies with the way this information is structured in language.

Other work from our lab has asked how people apprehend events with more complex internal structure. Linguistic theories propose that thematic roles follow a hierarchy which defines how these semantic representations are mapped onto syntax. For instance, according to an influential Thematic Hierarchy, Agents rank higher than Patients, which, in turn, rank higher than other thematic roles such as Goals, Sources, Instruments etc. (Baker, 1997; Jackendoff, 1990). In our own work (Wilson, Papafragou, Bunger, & Trueswell, 2011), we asked whether this hierarchy also characterises event apprehension. People were asked to watch static depictions of causative events involving Agents using an Instrument to direct a Patient into a Goal (e.g. a man using a rake to push leaves into a basket); their task was to identify event components by looking at them as quickly as possible and then pressing a response key. Participants were assigned to one of four conditions depending on the event component (or conceptual "thematic role") they were told to look at: specifically, they were instructed to "look at the person or animal who is performing the action" (Agent condition), "the object directly affected by the action" (Patient condition), "the goal or destination of the action" (Goal condition), or "the tool or body part used to make the action" (Instrument condition). Results showed that participants were faster at identifying an Agent in a causative event (as measured by eye movements to the corresponding region) compared to a Patient, Goal or Instrument, with Instruments being the slowest event components to be identified. Similarly, in a version of this experiment where participants were asked to describe the same scenes, Agents were mentioned most often in participants' descriptions, followed by Patients and Goals; Instruments were mentioned least often (Wilson et al., 2011). Together, these asymmetries confirm the Thematic Hierarchy and demonstrate clear affinities between the way people apprehend events from visual stimuli and the way they map these representations onto language. Furthermore, even though the above studies were all conducted with speakers of a single language (English), more recent extensions of the causative events paradigm have shown similar patterns in young learners of both English and Turkish, despite cross-linguistic differences in the surface realisation of

thematic roles in the two languages (Ünal, Trueswell, & Papafragou, 2017).

Mapping events onto language

How do speakers convert a non-linguistic event representation into a string of words produced sequentially? Eye-tracking studies of language production have pointed to a tight relationship between gaze and speech: speakers overwhelmingly fixate the objects in a display in the order in which they are going to talk about them (Altmann & Kamide, 2004; Bock et al., 2004). Since it is standardly assumed that shifts of visual attention are informative about the development of a message plan, eye-tracking methods can be useful in revealing how event concepts are mapped onto language. In the first study to explore this possibility, Griffin and Bock (2000) recorded speakers' direction of gaze as these speakers visually inspected and described static line drawings of simple actions (e.g. a woman shooting a man). Analysis of people's eye movements in relation to their linguistic choices led to the conclusion that there exists an initial rapid event apprehension stage that is temporally dissociable from any sentence generation stage. Later work has also found that message formulation can begin with the generation of a larger relational conceptual representation before linguistic encoding begins (e.g. Bock et al., 2004; Kuchinsky, Bock, & Irwin, 2011; Lee, Brown-Schmidt, & Watson, 2013).

Current evidence shows that the online planning of both messages and sentences is flexible and may operate over units of variable size. In one study, speakers were more likely to begin their descriptions of two-participant events (The dog chasing the mailman) with the participant that was perceptually more salient (i.e. cued with a subtle attentional cue), and mentioned the less salient participant later in the sentence, right after having directed their attention towards him (Gleitman, January, Nappa, & Trueswell, 2007). Thus linguistic formulation can proceed in small increments, guided by bottom-up visual information, without the need of extracting the full event structure. Other work shows that planning increments can be lexically sized. For instance, when describing simple objects (e.g. little house), speakers can encode the part of the message that expresses size (little) separately from the part of the message that identifies the object (house; Brown-Schmidt & Konopka, 2008). More recent evidence suggests that the size of the planning unit may depend on various factors such as the ease of apprehension of the relevant entities and relational structures within an event, the speakers' goals or processing constraints (Konopka & Meyer, 2014; van de Velde, Meyer, & Konopka, 2014). For instance, in one study where speakers were asked to describe two-participant events, they were more likely to prioritise a single event participant at the onset of linguistic formulation when the participant was easy to name but more likely to encode both characters early when the conceptual structure of the entire event was easy to extract (Konopka & Meyer, 2014).

Our own work has looked at how information is collected and put into words as people from different linguistic communities inspect events. We used as a point of departure the hypothesis that language-specific grammatical encoding biases may lead to systematized differences in the way that conceptual representations are mined for linguistic purposes (cf. Bock, 1995; Levelt, 1989; Slobin, 1996). To examine whether cross-linguistic differences might impact message planning, we conducted an eye tracking study in which adult speakers of Greek and English viewed a set of short animated motion events (e.g. a man skating to a snowman; Papafragou, Hulbert, & Trueswell, 2008). We chose these two languages because prior work had indicated that these languages represent two broad typological tendencies in the encoding of motion (Papafragou, Massey, & Gleitman, 2002; Slobin, 1996; Talmy, 1985). In English, information about manner of motion is typically encoded in the main verb (e.g. *A man skated* …) and path information is mentioned later, usually in a post-verbal prepositional phrase (… *to the snowman*). In Greek, by contrast, information about path is usually encoded in the main verb (e.g. *Enas andras pige sto hionanthropo* "A man went to-the snowman … ") and manner information is mentioned later, usually in a post-verbal prepositional phrase (*me patinia* "on skates") or omitted altogether. The visual stimuli were constructed so that manner and path information was depicted in regions spatially separated from each other: manners corresponded to the vehicle that propelled the moving agent (e.g. the skates in the earlier example) and paths corresponded to the object that served as the endpoint of the motion trajectory (e.g. the snowman). Events unfolded for 3 s, at which point a beep was heard, and the last clip from the event remained on the screen for another 2 s. Half of the participants were asked to describe the events after hearing the beep (Linguistic task). The other half were asked to watch the events carefully in preparation for a memory test (Non-Linguistic task).

As we expected, there were cross-linguistic differences in the way participants described motion events in the Linguistic task: For English speakers, 78% of all sentences contained a manner verb, as compared to only

32% for Greek speakers. In addition, there were cross-linguistic differences in the patterns of attention to those events as participants prepared to describe them, with speakers of each language turning their attention very early to those event components (manner or path) that they planned to encode in the verb of their event description. Specifically, Greek speakers were more likely than English speakers to fixate the path endpoint first (e.g. the snowman) rather than the manner of motion region (e.g. the skates). After about a second and a half, Greek speakers turned their attention to manner, while English speakers focused on the path endpoint, presumably as a result of the preparation of relevant post-verbal modifiers in each language. These eye movement patterns were repeated after the beep while people were describing aloud the events. This pattern is in accord with single-language eye movement production studies, where participants' preparation to describe regions of a scene was preceded by fixations on these regions (e.g. Altmann & Kamide, 2004; Griffin & Bock, 2000). Importantly, it is the first evidence that – as anticipated in *Speaking* – event apprehension during production differs cross-linguistically: where languages differ from each other in how they encode events, this difference shows up in how events are interrogated during speech planning.

In the Non-Linguistic task, where people were asked to simply study (but not describe) the events, attention allocation as the events unfolded was strikingly similar for both language groups: overall, people prioritised looks to the path endpoint and inspected the manner of motion slightly later. However, there was one striking cross-linguistic difference in the Non-Linguistic task: late in each trial, after the event had unfolded and its last frame remained on the screen, English speakers spent more time inspecting the path endpoint (e.g. the snowman) rather than the manner of motion (e.g. the skates) as compared to the Greek speakers who tended toward the opposite pattern. Later studies showed that this finding was not due to deep effects of language on attention but rather the result of the online use of language to encode details of the events that participants might forget: when participants performed a similar Non-Linguistic task under linguistic interference (repeating back numbers), this effect disappeared, but when participants engaged in the task under nonlinguistic interference (tapping a rhythm), the effect persisted (Trueswell & Papafragou, 2010). Together these findings from various versions of the Non-Linguistic task support the conclusion that event perception is independent of the viewer's native language.

These findings illustrate that language production relies on language-specific representations of event structure that are not assembled in situations that do not involve linguistic communication (e.g. during free perception of events). Further work has confirmed that preparing to describe motion events leads to distinct shifts of attention (compared to non-linguistic tasks such as free event viewing) already in 3- to 4-year old children (Bunger, Trueswell, & Papafragou, 2012), and generalises to young learners across language groups (Bunger, Skordos, Trueswell, & Papafragou, 2018; Bunger, Skordos, Trueswell, & Papafragou, 2016; for further cross-linguistic approaches to message formulation, see Brown-Schmidt & Konopka, 2008; Norcliffe, Konopka, Brown, & Levinson, 2015).

In a subsequent study we began to explore in greater detail how adult speakers take the conceptual structure of an event into account as they formulate an utterance to describe it. We built on the observation that speakers tend to repeat linguistic structures that they have recently used or observed others using (*structural priming*; Bock, 1986; Levelt & Kelter, 1982; see Pickering & Ferreira, 2008, for an overview). Our reasoning was that the kind of abstract structures that speakers tend to repeat in a priming paradigm can offer evidence for the nature of the conceptual motion event representations that are accessed during language production.

In this study, English speakers read prime sentences with motion meanings and later had to describe target motion events (Bunger, Papafragou, & Trueswell, 2013). We varied the degree of prime-target overlap, such that, for a given target (e.g. an event in which an alien drove a car into a cave), the prime would (a) overlap with the target in terms of both broad event type (i.e. motion defined by a path) and specific verb that could be used to describe the event (*The zebra on the motorcycle entered the garage*); (b) overlap with the target in terms of event type but not specific verb (*The man in the helicopter circled the tower*); (c) have no overlap with the target (*The nurse with the freckles baked a pie*). We found that speakers were more likely to mention the path of a target motion event if they had been primed with a sentence that overlapped in event type, even when prime sentences did not provide a specific verb that could be re-used to describe the target event. In a subsequent experiment, the effect held even when the order in which motion event components were evoked in prime sentences was switched from manner before path (e.g. *The zebra on the motorcycle entered the garage*) to the less-canonical for English path before manner (*The zebra entered the garage on the motorcycle*). Together, this set of findings offers evidence for a kind of conceptual priming, demonstrating that activation of event structure has implications for message planning: priming speakers with particular

event structures increased the accessibility of those event types, and influenced, in turn, the event information that speakers selected to talk about (path/manner of motion). Our findings lead to the expectation that such priming should also occur when the primes are not sentences but non-linguistic stimuli carrying similar types of information.

Two additional findings from this study are worth highlighting. First, priming had implications not only for the information that speakers chose to communicate about target motion events but also for the way that information was mapped onto linguistic elements: when prime sentences both presented verbs that could be re-used to describe target events *and* presented motion event components in the canonical manner-before-path order, speakers were primed to encode information about the manners and paths of target motion events *in the same locations that they were presented in prime sentences*: path information in main verbs, and manner information in subjects (as in the earlier prime example *The zebra on the motorcycle entered the garage*). This happened despite the fact that English typically encodes manner, and not path, information in motion verbs. However, priming of motion event component encoding was not seen when prime sentences presented speakers with information about motion event components in the non-canonical path-before-manner order. Thus speakers were willing to violate their bias to encode manner information in verbs, but not their preference to present manner information before path information (rooted in the syntactic tendency in English to encode manner information in verbs and path information in post-verbal modifiers; cf. Talmy, 1985, and earlier discussion). As a result, conceptual priming failed when pitted against language-specific biases about the order in which information about a motion event should be encoded.

Second, event structure priming led to an increase in the use of primed syntactic frames in descriptions of target events. Priming of syntactic structure in this study was only successful when primes provided both conceptual (event type) and lexical (specific verb) overlap with target events (even though throughout these experiments we observed priming of syntactic structure that was independent of the repetition of primed verbs). For present purposes, this result bolsters the conclusion that message content and structure may themselves be influenced by conceptual information in the input.

The work reviewed in the previous section demonstrated that event components such as agents and patients are extracted quickly and spontaneously even from very brief demonstrations of events. The research reviewed here suggests that conceptual event components such as manner and path are not only readily extracted from dynamic motion events but also made available for packaging into a linguistic form.

Generating event messages in conversation

So far we have considered single, simple descriptions of event content. However, in both cognition and language, event encoding may involve multiple levels of abstraction and degrees of specificity. For example, the same event can be described as *A man and a woman biking up the hill*, *A couple biking*, or simply *Biking*. In everyday conversation, the decision about which event components to mention and which to omit is the result of a tradeoff between two types of pressures: the pressure to construct a message that is successful (i.e. as informative and relevant as expected in a given communicative exchange; Grice, 1975; Sperber & Wilson, 1986) and the pressure to construct a message that is efficient (i.e. easily recoverable by the listener with the least effort for the speaker; cf. Levinson, 1995; Shannon, 1948; Zipf, 1949). To construct successful messages, speakers need to take into account the knowledge and beliefs of their listeners (Clark & Marshall, 1981; Levelt, 1989), by adjusting their speech to what their listeners can see, what information was shared in prior discourse, and other kinds of common ground (Brennan & Clark, 1996; Galati & Brennan, 2010; Heller, Gorman, & Tanenhaus, 2012; Nadig & Sedivy, 2002). To construct efficient messages, speakers can omit or abbreviate already known or easily inferable information (Ariel, 1990; Bard et al., 2000; Fowler & Housum, 1987; Samuel & Troicki, 1998; Galati & Brennan, 2010), as the linguistic system relies on listeners' ability to draw inferences and enrich the selective linguistic encoding of information (Grice, 1975; Sperber & Wilson, 1986). These pragmatic considerations in message planning have been extensively studied in the context of referential disambiguation where a speaker has to specify one of many potential referents in a display using appropriate referential devices (e.g. the brown cup/the cup/it; for early instantiations, see Krauss & Glucksberg, 1977; Krauss & Weinheimer, 1964). Fewer studies have examined how speakers select information to be encoded in event description, which is characterised by less restricted conversational goals than disambiguation and where, as a result, the interplay among conceptual representation, pragmatic considerations and language-specific pressures in message planning is more intense.

Existing experimental evidence suggests that speakers adjust their event descriptions in anticipation of

what listeners would find easy to infer. In a seminal study, Brown and Dell (1987) found that, in retelling stories involving instrument events, speakers tended to omit typical/easily predictable instruments (e.g. stabbing with a knife) but frequently mentioned atypical/highly unpredictable instruments (e.g. stabbing with an icepick), anticipating that their listeners would not be able to recover the unpredictable event component. The authors described such adaptations as *generic* adjustments, since instrument typicality is a property accessible to any member of the linguistic community who shares common knowledge about how various events are performed.

Do speakers also take into account the knowledge and beliefs of their *specific* interlocutor during speech planning? In the Brown and Dell (1987) study, whether the listener could see the events or not had no effect on whether the speaker mentioned the instrument of an action – a finding that seemed to suggest that listeners do not design their utterances with their specific addressee's mental state in mind. A modified version of this study by Lockridge and Brennan (2002), however, did find effects of the listener's visual perspective: speakers were more likely to include atypical instruments in the story when their listener did not have access to pictures depicting the events compared to when listeners could see these pictures. Furthermore, these production adjustments happened early in the clause in a way that suggested early sensitivity to listener knowledge during message planning. Lockridge and Brennan attributed the discrepancy in the findings to the fact that, in Brown and Dell's study, the listener was a confederate of the experimenter, but in their design, the listener was another naïve participant. Other work has corroborated the idea that the listener's involvement in a task and the speaker's assumptions about this involvement affect speech planning decisions (see Brennan, Galati, & Kuhlen, 2010; Kuhlen & Brennan, 2010; Schober, 1993).

In recent work, we have directly tested the role of the listener's communicative profile in adults' and children's production choices (Grigoroglou & Papafragou, 2018). Unlike past studies that used story retelling, we asked 4- to 5-year-olds and adults to watch videotaped events involving an agent performing everyday actions with typical and atypical instruments (e.g. a man digging a hole with a shovel/a plate) and to spontaneously describe them to listeners who could either see or not see the events. The profile of the listener was manipulated across three experiments: the listener was either silent and without a clear goal, silent but with a stated goal (i.e. to draw pictures of the events described), or interactive with the same stated goal. Results showed that both adults and 4- to 5-year-olds were more likely to add information about instruments when communicating with an interactive listener (as opposed to a silent one, regardless of whether the silent listener had a goal). Furthermore, both groups were more likely to mention atypical instruments in their descriptions compared to typical instruments. However, only adults adjusted their speech to the visual perspective of the listener, mentioning instruments more often when the addressee could not see them. Children failed to do so and their descriptions remained, overall, under-informative (see also Grigoroglou & Papafragou, 2019, for similar results in a task closer to standard referential communication paradigms).

These results suggest that adults' message planning decisions are affected by pragmatic considerations about the listener's informational needs, especially when these needs arise from interactions with the addressee and are assumed to be genuine. Furthermore, adults incorporate both generic (typicality-based) and specific (visual-access-based) adaptations to their addressees in free production (cf. also Lockridge & Brennan, 2002). Additionally, however, these findings support a nuanced model of listener-oriented adaptations: for children, maintaining a "particular" model of the listener in some cases seems to be associated with additional costs compared to a more "generic" one, a finding reminiscent of Brown and Dell (1987). In our study, the costs seem to emanate from the need to constantly track the listener's visual perspective when that perspective conflicts with one's own. Echoes of this phenomenon exist even in our adult data (as well as those in Lockridge & Brennan, 2002), where the effect of visual access was smaller than the effect of typicality. It follows that the ability to adjust one's speech to the perspective of the addressee should not viewed as an "all or nothing" ability but rather as a set of distinct cognitive abilities along a continuum (see Liebal, Carpenter, & Tomasello, 2013; Moll & Kadipasaoglu, 2013; for related views). This nuanced perspective on developmental findings is consistent with evidence from adult psycholinguistics suggesting that the process of designing messages with a particular addressee's needs in mind depends on several cognitive factors (e.g. cognitive load, saliency of privileged information), relies on representations of different types of common ground in memory that are independent from each other (Brown-Schmidt & Duff, 2016), and can fail under time pressure or more complex perspective-tracking demands (see Konopka & Brown-Schmidt, 2014, for a review).

The pragmatic considerations outlined in this section are presumed to reflect universal aspects of the architecture of human communication, such as the desire to offer

informative messages (Grice, 1975; Sperber & Wilson, 1986). However, their specific implementation during message planning is likely to vary cross-linguistically. As Levelt (1989) observed, not all aspects of preverbal messages need to proceed from elaborations of the communicative intention of the speaker but may simply reflect language-specific grammatical requirements (such as the obligatory inclusion of tense or aspect), and if so, should be expected to differ across languages. This expectation is borne out. Extending Brown and Dell (1987), Papafragou, Massey, and Gleitman (2006) asked 8-year-old children and adult speakers of English and Greek to describe motion events involving typical or atypical manners (a man walking vs. running up the stairs). Greek speakers in both age groups were more likely to encode manner of motion when it was atypical (e.g. *Enas andras aneveni tis skales trehontas* "a man is-ascending the stairs running") compared to when it was typical and thus easily inferable by any listener even if unmentioned (e.g. *Enas andras aneveni tis skales* "a man is-ascending the stairs" [walking]). English speakers, by contrast, encoded manner regardless of typicality (e.g. *A man ran up the stairs/walked up the stairs*). The difference can be explained by the fact that, as alluded to already, Greek tends to encode path information in verbs and manner information in optional linguistic constituents (e.g. post-verbal modifiers), whereas English strongly prefers to encode manner of motion in the main verb. Because of this strong encoding preference, manner of motion is selected for mention during message generation in English regardless of its communicative value.

Conclusions

Speaking is one of the most complex skills available to humans. Its origins are hidden in each speaker's private thoughts, and yet the act of speaking is public and tightly coupled with the dynamics of human interaction. Furthermore, speaking is unique to humans and emerges early in life, but its outputs are vastly different from one linguistic community to the next. Pim Levelt's seminal book on this topic, published 30 years ago, synthesised these perspectives and dissected the multiple processes that allow humans to transform intentions and thoughts into fluent utterances. A considerable part of the book was dedicated to how speakers begin planning their speech by conceptualising a preverbal content that will form the input to later processes of formulation and articulation. Here we have sketched a programme of research that tackled many of the foundational proposals about conceptualisation laid out in *Speaking* that have continued to both fascinate and elude researchers since the book's publication.

This research points to several key findings about the workings of the conceptualisation stage in language production. First, evidence from how viewers process event types and event roles supports the position that language production builds on underlying conceptual representations and that both mature and young language users use such representations to structure incoming experience. Second, as experiments on motion events show, conceptual representations are mobilised on-line when speakers plan their utterances in a way consistent with language-specific event encoding. This occurs even though conceptual representations assembled during free viewing of the events in preparation of a non-linguistic task are unaffected by language-specific encoding requirements. Furthermore, conceptual representations about event structure can be activated through priming and this activation affects the selection of event components during message planning (as well as downstream processes of utterance formulation). Third, message planning is also affected by pragmatic considerations that employ models of the listener at various levels of specificity (including, e.g. both what a specific listener knows in a given context and what any generic listener would know under normal circumstances). Additionally, the pragmatic attunement of message generation interacts with language-specific factors, such that preverbal message-encoding is different in speakers of different languages. Together, these novel findings build on and further articulate several of the broad hypotheses about the role of conceptualisation proposed in *Speaking*. They also open up several questions for further work on how speakers conceptualise events (and other to-be-mentioned entities) in preparation for language production.

One major question is how findings from the thematic structure of event representations extend to other thematic roles (see Lakusta & Carey, 2014; Lakusta & Landau, 2005, 2012; Papafragou, 2010 on the status of Sources vs. Goals of motion in language and cognition) and, more generally, other abstract aspects of preverbal messages. For instance, beyond thematic structure, preverbal messages have been hypothesised to reflect the abstract temporal profile of events – for instance, whether an event has an inherent boundary or not (cf. the contrast in the events encoded by the sentences *Mary ate a snack* vs. *Mary ate snacks* respectively; Jackendoff, 1990). Recent work from our lab has shown that both adults and young learners are sensitive to event boundedness distinctions in a range of non-linguistic tasks, lending support to the idea that temporal event structure is part of preverbal message content (Ji & Papafragou, 2017, 2018). Notice that, for all these

phenomena, evidence for homologies between conceptual and linguistic representations needs to be confirmed across languages. This is particularly pressing since such homologies (mostly posited on the basis of evidence from English speakers) leave open the question whether typological differences across languages might affect non-linguistic apprehension of events. Our own investigations from the domain of motion have revealed core similarities in event perception despite cross-linguistic differences (Papafragou et al., 2002; Papafragou et al., 2008; cf. also Ünal et al., 2017). However, there is currently a lively discussion of how concepts and language come together across linguistic communities (see Malt & Wolff, 2010; Ünal & Papafragou, 2016 for recent reviews).

A second major issue concerns how non-linguistic event representations are fed into the processes that formulate utterances. *Speaking* leaves open the possibility that chunks of preverbal messages are converted into utterances in an incremental fashion but at present several theoretical options remain open about the scope of planning including both the factors that affect the size of the planning unit (Bock, Irwin, Davidson, & Levelt, 2003; Bock et al., 2004; Brown-Schmidt & Konopka, 2008; and earlier discussion) and the way information is linearised into sequences of units (Norcliffe et al., 2015; see Konopka & Brown-Schmidt, 2014, for a review). More research is needed on how transitions from messages to the formulation stage proceed, and on the timing and interaction of these processes.

Finally, the findings presented here provide evidence for event conceptualisation as a distinct level of representation that organises the continuous flux of visual information into meaningful units which can then be mapped onto language. Although, for the most part, event conceptualisation has been studied within single events, pre-vetted by researchers in the laboratory, event apprehension does not happen in a void but is embedded into broader, often fairly complex contexts. The work we reviewed in the last section demonstrates how contextual factors such as the communicative status of the listener affect message planning. It remains to be seen whether context also affects event conceptualisation. For instance, a simple event such as making the bed can mean different things depending on the context in which it occurs: the event endpoint (i.e. whether the bed is made) may vary depending on whether the perceiver is the mother of a teenager or the director of a 5-star hotel. Future research could explore how higher-order knowledge (involving different kinds of information about social norms, people's histories and intentions, the perceiver's mood) affects the way people conceptualise events and plan their messages.

Final thoughts

At its most general, the research presented here has tried to contribute to our understanding of how the process of conceptualising the world exerts an influence on the process of linguistically describing it. Because conceptualisation is the backbone of language production, its inner workings have traditionally been the subject of foundational but hard-to-test hypotheses that have characterised psycholinguistics for decades. Here we have argued that, through a combination of online behavioural, cross-linguistic and developmental methods, these classic ideas can be experimentally investigated in ways that can turn conceptualisation from a fascinating mystery into an empirically tractable problem.

Acknowledgement

Preparation of this article was supported by NSF grant #1632849 (A.P.). M.G. would like to acknowledge support from a Doctoral Fellowship from the University of Delaware. Both authors would like to thank Pim Levelt for inspiration and discussion, and the organisers and audience of IWLP 10 at the Max Planck Institute for Psycholinguistics at Nijmegen for helpful input.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by NSF Grant BCS-1632849 to Anna Papafragou.

References

Altmann, G., & Kamide, Y. (2004). Now you see it, now you don't: Mediating the mapping between language and the visual world. In J. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action* (pp. 347–386). New York, NY: Psychology Press.

Ariel, M. (1990). *Accessing noun-phrase antecedents*. London: Routledge.

Baker, M. C. (1997). Thematic roles and syntactic structure. In L. Haegeman (Ed.), *Handbook of generative syntax* (pp. 73–137). Dordrecht: Kluwer.

Bard, Ellen Gurman, Anderson, Anne H., Sotillo, Catherine, Aylett, Matthew, Doherty-Sneddon, Gwyneth, & Newlands, Alison. (2000). Controlling the Intelligibility of referring expressions in dialogue. *Journal of Memory and Language*, *42*(1), 1–22. doi:10.1006/jmla.1999.2667

Bock, J. K. (1986). Meaning, sound, and syntax: Lexical priming in sentence production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 575–586.

Bock, J. K. (1995). Sentence production: From mind to mouth. In J. Miller & P. Eimas (Eds.), *Handbook of perception and*

cognition: Speech, language, and communication (Vol. 11, pp. 181–216). New York: Academic Press.

Bock, K., Irwin, D., & Davidson, D. (2004). Putting first things first. In J. Henderson & F. Ferreira (Eds.), *The interface between language, vision and action: Eye movements and the visual world* (pp. 249–317). New York and Hove: Psychology Press.

Bock, Kathryn, Irwin, David E., Davidson, Douglas J., & Levelt, W. J. M. (2003). Minding the clock. *Journal of Memory and Language*, *48*(4), 653–685. doi:10.1016/S0749-596X

Bock, J. K., & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 945–984). San Diego, CA: Academic Press.

Bowerman, M., & Choi, S. (2001). Shaping meanings for language: Universal and language-specific in the acquisition of spatial semantic categories. In M. Bowerman & S. Levinson (Eds.), *Language acquisition and conceptual development* (pp. 475–511). Cambridge: Cambridge University Press.

Bowerman, M., & Levinson, S. C. (2001). *Language acquisition and conceptual development*. Cambridge: Cambridge University Press.

Brennan, S. E., & Clark, H. H. (1996). Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology*, *22*(6), 1482–1493.

Brennan, S. E., Galati, A., & Kuhlen, A. K. (2010). Two minds, one dialog: Coordinating speaking and understanding. *Psychology of Learning and Motivation*, *53*, 301–344.

Brown, P., & Dell, G. (1987). Adapting production to comprehension: The explicit mention of instruments. *Cognitive Psychology*, *19*, 441–472.

Brown-Schmidt, S., & Duff, M. C. (2016). Memory and common ground processes in language use. *Topics in Cognitive Science*, *8*(4), 722–736.

Brown-Schmidt, S., & Konopka, A. E. (2008). Little houses and casas pequeñas: Message formulation and syntactic form in unscripted speech with speakers of English and Spanish. *Cognition*, *109*, 274–280.

Bunger, A., Papafragou, A., & Trueswell, J. C. (2013). Event structure influences language production: Evidence from structural priming in motion event description. *Journal of Memory and Language*, *69*(3), 299–323.

Bunger, A., Skordos, D., Trueswell, J., & Papafragou, A. (2018). *How children attend to events before speaking: Cross-linguistic evidence from the motion domain*. Manuscript submitted for publication.

Bunger, A., Skordos, D., Trueswell, J. C., & Papafragou, A. (2016). How children and adults encode causative events cross-linguistically: Implications for language production and attention. *Language, Cognition and Neuroscience*, *31*(8), 1015–1037.

Bunger, A., Trueswell, J. C., & Papafragou, A. (2012). The relation between event apprehension and utterance formulation in children: Evidence from linguistic omissions. *Cognition*, *122* (2), 135–149.

Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.

Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. H. Joshi, B. I. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge: Cambridge University Press.

Dobel, C., Glanemann, R., Kreysa, H., Zwitserlood, P., & Eisenbeiss, S. (2010). Visual encoding of coherent and non-coherent scenes. In E. Pedersen & J. Bohnemeyer (Eds.), *Event representation in language: Encoding events at the language cognition interface* (pp. 189–215). Cambridge: Cambridge University Press.

Dobel, C., Gumnior, H., Bölte, J., & Zwitserlood, P. (2007). Describing scenes hardly seen. *Acta Psychologica*, *125*, 129–143.

Dowty, D. (1991). Thematic proto-roles and argument selection. *Language*, *67*, 547–619.

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

Galati, A., & Brennan, S. E. (2010). Attenuating information in spoken communication: For the speaker, or for the addressee? *Journal of Memory and Language*, *62*(1), 35–51.

Gentner, D., & Goldin-Meadow, S. (Eds.). (2003). *Language in mind: Advances in the study of language and thought*. Cambridge, MA: MIT Press.

Gleitman, L. R., January, D., Nappa, R., & Trueswell, J. C. (2007). On the give and take between event apprehension and utterance formulation. *Journal of Memory and Language*, *57*, 544–569.

Gleitman, L., & Papafragou, A. (2016). New perspectives on language and thought. In K. Holyoak & R. Morrison (Eds.), *Oxford handbook of thinking and reasoning* (2nd , pp. 543–568). New York, NY: Oxford University Press. doi:10.1093/oxfordhb/9780199734689.013.0028

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and semantics: Speech acts* (Vol. 3, pp. 41–58). New York, NY: Academic Press.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psychological Science*, *11*(4), 274–279.

Grigoroglou, M., & Papafragou, A. (2018, July). *Children's adjustments to listener needs in spontaneous event descriptions*. Poster presented at the International Workshop on Language Production (IWLP), Nijmegen.

Grigoroglou, M., & Papafragou, A. (2019). Interactive contexts increase informativeness in children's referential communication. *Developmental Psychology*. Advanced online publication. doi:10.1037/dev0000693.

Hafri, A., Papafragou, A., & Trueswell, J. C. (2013). Getting the gist of events: Recognition of two-participant actions from brief displays. *Journal of Experimental Psychology: General*, *142*(3), 880–905.

Hafri, A., Trueswell, J. C., & Strickland, B. (2018). Encoding of event roles from visual scenes is rapid, spontaneous, and interacts with higher-level visual processing. *Cognition*, *175*, 36–52.

Heller, D., Gorman, K. S., & Tanenhaus, M. K. (2012). To name or to describe: Shared knowledge affects referential form. *Topics in Cognitive Science*, *4*, 290–305.

Jackendoff, R. (1990). *Semantic structures*. Cambridge, MA: MIT Press.

Ji, Y., & Papafragou, A. (2017). Viewers' sensitivity to abstract event structure. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th annual conference of the cognitive science society* (pp. 594–599). Austin, TX: Cognitive Science Society.

Ji, Y., & Papafragou, A. (2018). Midpoints and endpoints in event perception. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings of the 40th annual conference of the cognitive science society* (pp. 1877–1882). Austin, TX: Cognitive Science Society.

Konopka, A. E., & Brown-Schmidt, S. (2014). Message encoding. In V. Ferreira, M. Goldrick, & M. Miozzo (Eds.), *The Oxford handbook of language production* (pp. 3–20). Oxford: Oxford University Press.

Konopka, A. E., & Meyer, A. S. (2014). Priming sentence planning. *Cognitive Psychology*, *73*, 1–40.

Krauss, R. M., & Glucksberg, S. (1977). Social and nonsocial speech. *Scientific American*, *236*, 100–105.

Krauss, R. M., & Weinheimer, S. (1964). Changes in reference phrases as a function of frequency of usage in social interaction: A preliminary study. *Psychonomic Science*, *1*, 113–114.

Kuchinsky, S. E., Bock, K., & Irwin, D. E. (2011). Reversing the hands of time: Changing the mapping from seeing to saying. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 748–756.

Kuhlen, A. K., & Brennan, S. E. (2010). Anticipating distracted addressees: How speakers' expectations and addressees' feedback influence storytelling. *Discourse Processes*, *47*(7), 567–587.

Lakusta, Laura, & Carey, Susan. (2014). Twelve-month-old infants' encoding of goal and source paths in agentive and non-agentive motion events. *Language Learning and Development*, *11*(2), 152–175. doi:10.1080/15475441.2014.896168

Lakusta, Laura, & Landau, Barbara. (2005). Starting at the end: the importance of goals in spatial language. *Cognition*, *96* (1), 1–33. doi:10.1016/j.cognition.2004.03.009

Lakusta, Laura, & Landau, Barbara. (2012). Language and memory for motion events: Origins of the asymmetry between source and goal paths. *Cognitive Science*, *36*(3), 517–544. doi:10.1111/cogs.2012.36.issue-3

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–136). New York: Wiley.

Lee, E. K., Brown-Schmidt, S., & Watson, D. W. (2013). Ways of looking ahead: Incrementality in language production. *Cognition*, *129*, 544–562.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: The MIT Press.

Levelt, W. J. M., & Kelter, S. (1982). Surface form and memory in question answering. *Cognitive Psychology*, *14*, 78–106.

Levin, B., & Rappaport-Hovav, M. (2005). *Argument realization*. Cambridge: Cambridge University Press.

Levinson, S. C. (1995). Three levels of meaning: Essays in honor of Sir John Lyons. In F. R. Palmer (Ed.) *Grammar and meaning* (pp. 90–115). Cambridge: Cambridge University Press.

Liebal, K., Carpenter, M., & Tomasello, M. (2013). Young children's understanding of cultural com- mon ground. *British Journal of Developmental Psychology*, *31*(1), 88–96.

Lockridge, C. B., & Brennan, S. E. (2002). Addressees' needs influence speakers' early syntactic choices. *Psychonomic Bulletin & Review*, *9*(3), 550–557.

Majid, A., Boster, J. S., & Bowerman, M. (2008). The cross-linguistic categorization of everyday events: A study of cutting and breaking. *Cognition*, *109*, 235–250.

Malt, B. C., & Wolff, P. M. (Eds.). (2010). *Words and the mind: How words capture human experience*. Oxford: Oxford University Press.

Moll, Henrike, & Kadipasaoglu, Derya. (2013). The primacy of social over visual perspective-taking. *Frontiers in Human Neuroscience*, *7*. doi:10.3389/fnhum.2013.00558

Nadig, A. S., & Sedivy, J. C. (2002). Evidence of perspective-taking constraints in children's on-line reference resolution. *Psychological Science*, *13*(4), 329–336.

Norcliffe, E., Konopka, A. E., Brown, P., & Levinson, S. C. (2015). Word order affects the time course of sentence formulation in Tzeltal. *Language, Cognition and Neuroscience*, *30*(9), 1187–1208.

Papafragou, Anna. (2010). Source-goal asymmetries in motion representation: Implications for language production and comprehension. *Cognitive Science*, *34*(6), 1064–1092. doi:10.1111/

Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, *108*, 155–184.

Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, 'n' roll: The representation of motion in language and cognition. *Cognition*, *84*, 189–219.

Papafragou, A., Massey, C., & Gleitman, L. (2006). When English proposes what Greek presupposes: The cross-linguistic encoding of motion events. *Cognition*, *98*(3), B75–B87.

Paul, H. (1970). The sentence as the expression of the combination of several ideas. In A. L. Numenthal (Trans.), *Language and psychology: Historical aspects of psycholinguistics* (pp. 20–31). New York: Wiley (Original work published in 1886).

Pickering, M. J., & Ferreira, V. S. (2008). Structural priming: A critical review. *Psychological Bulletin*, *134*, 427–459.

Pinker, S. (1989). Learnability and cognition: The acquisition of argument structure. In *Language* (Vol. 68, pp. xiv, 411). Retrieved from http://www.isrl.uiuc.edu/~amag/ langev/paper/pinker89book.html

Radvansky, G. A., & Zacks, J. M. (2014). *Event cognition*. New York: Oxford University Press.

Samuel, S. G., & Troicki, M. (1998). Articulation quality is inversely related to redundancy when children or adults have verbal control. *Journal of Memory and Language*, *39*, 175–194.

Schober, M. F. (1993). Spatial perspective-taking in conversation. *Cognition*, *47*, 1–24.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*, 379–423.

Shipley, T. F., & Zacks, J. M. (Eds.). (2008). *Understanding events: From perception to action*. New York, NY: Oxford University Press.

Slobin, D. (1996). From 'thought and language' to 'thinking for speaking'. In J. Gumperz & S. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). New York, NY: Cambridge University Press.

Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition* (2nd ed.). Harvard: Harvard University Press.

Talmy, L. (1985). Lexicalization patterns: Semantic structure in lexical forms. In T. Shopen (Ed.), *Language typology and syntactic description* (pp. 57–149). New York, NY: Cambridge University Press.

Trueswell, J. C., & Papafragou, A. (2010). Perceiving and remembering events cross-linguistically: Evidence from dual-task paradigms. *Journal of Memory and Language*, *63*(1), 64–82.

Ünal, E., & Papafragou, A. (2016). Interactions between language and mental representations. *Language Learning*, *66*(3), 554–580.

Ünal, E., Trueswell, J., & Papafragou A. (2017, July). *How children map event participants onto language*. Talk presented at Symposium

on 'Encoding events in language and cognition', 14th International Congress for the Study of Child Language, Lyon.

van de Velde, M., Meyer, A., & Konopka, A. E. (2014). Message formulation and structural assembly: Describing "easy" and "hard" events with preferred and dispreferred syntactic structures. *Journal of Memory and Language*, *71*, 124–144.

Webb, A., Knott, A., & MacAskill, M. R. (2010). Eye movements during transitive action observation have sequential structure. *Acta Psychologica*, *133*, 51–56.

Wilson, F., Papafragou, A., Bunger, A., & Trueswell, J. (2011). *Rapid extraction of event participants in caused motion events*. Proceedings from the 33rd Annual Meeting of the Cognitive Science Society, Erlbaum, Hillsdale, NJ.

Wundt, W. (1970). The psychology of the sentence. In A. L. Blumenthal (Trans.), *Language and psychology: Historical aspects of psycholinguistics* (pp. 20–31) (Original work published in 1900). New York, NY: Wiley.

Zipf, G. K. (1949). *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley Press.

Zwitserlood, P., Bölte, J., Hofmann, R., Meier, C. C., Dobel, C., & Bolhuis, J. J. (2018). Seeing for speaking: Semantic and lexical information provided by briefly presented, naturalistic action scenes. *PLOS ONE*, *13*(4), e0194762.