Events in Language and Cognition - 2016

Organizers

Eva Wittenberg, UC San Diego
Melissa Kline, MIT
Joshua Hartshorne, Boston College

Program Committee

Ben Ambridge, Curt Anderson, Ryan Blything, Peter Culicover, Annelot de Rechteren van Hemert, Monique Flecken, Nate George, Tilbe Goksun, Roberta Golinkoff, Angela He, Anita Krishnan, Laura Lakusta, Kyle Mahowald, Titus von der Malsburg, Emily Morgan, Mark Myslin, Ken Nakatami, Laura Niemi, Liuba Papeo, Shannon Pruden, Jason Rothman, Elliot Saltzman, David Townsend, Ercenur Unal.

Helpers

Edith Kaan, Marc Matthews, Suhas Arehalli, Michelle Waters, George Collins

Workshop Description

Understanding how speakers wrap event conceptualizations into linguistic descriptions is crucial for both linguistic theory and psychology. A number of rich linguistic theories have been proposed to account for the observed ways that meaning maps to syntax within and across languages, but their psychological status remains unclear. These theories often propose specific representational architecture, ranging from prototype theories to predicate decomposition. How are these conceptual models and mappings grounded in nonlinguistic cognition? On the side of cognitive science, our understanding of event representation, especially in infancy, has advanced dramatically in the past several decades, potentially opening up new possibilities for evaluating the plausibility of proposed argument structure theories. What can the understanding of event perception and cognition teach us about the nature of semantic representations for language, and how can psycholinguistic evidence contribute to research on event structure?

We aim to bring together psycholinguists interested in how language maps onto event structure, adding momentum to a growing field of interdisciplinary scholarship of event representation in linguistics, psycholinguistics, and cognitive science. We define three broad questions of interest:

(1) How do children and adults mentally represent events at the level of abstraction that is encoded in language?

(2) How is the structure of events represented in and conveyed by language?

(3) How to the constraints of the infant cognitive system and questions of learnability constrain our theories of language and how does language constrain our theories of the infant cognitive system?

Several fields are currently approaching each of these questions, though with different theoretical underpinnings, methods, and terminologies. Given that, what should we be working on as an interdisciplinary field? Are there specific empirical questions that would be relevant to all theoretical perspectives? And what is necessary in order to bring together psycholinguists working on different questions? In this workshop, we want to both foster an exchange of recent work, and set a possible agenda for psycholinguistic research on event structure as conveyed by language.

Workshop Schedule

8:30 Welcome, poster previews

9:00 **Elsi Kaiser**, Keynote Speaker

9:45 **Gerwien & von Stutterheim**
Grammatical constraints on event packaging and potential effects on the segmentation of the perceptual stream.

10:15 **Lundquist, Corley, Ramchand, Sorace, & Tungseth**
Crosslinguistic variation in event conceptualization: Evidence from the causative alternation.

10:45 Break

11:15 **Colicover, Alishahi, & Vaiksnoraite**
The constructional evolution of grammatical functions and the thematic representation of events.

12:15 **Altmann**
Multiple object-states represent events: A new account of event representation.

12:45 Lunch Break

1:45 **Jesse Snedeker**, Keynote Speaker

2:30 **Foppolo, Panzeri, Greco, & Carminatti**
The incremental processing of accomplishment predicates.

3:00 **Venhuizen, Brouwer, & Crocker**
When the food arrives before the menu: Modeling event-driven surprisal in language comprehension.

3:30 Break

3:45 Poster Session

4:45 **Jeremy Skipper & Elliot Saltzmann** (With Kaiser and Snedeker)
Panel Discussion

Abstracts

Multiple Object-states Represent Events: A new account of event representation

Events cause change. Even simple organisms like amoebae[1] react to change (e.g. Westendorf, Negrete, Bae, Sandmann, Bodenschatz, & Beta, 2013). Our own (human) ability to notice, track, represent, recall, and communicate change is at the heart of human function; from our most peripheral sensory systems to the highest levels of cognitive representation and processing. Here, we consider the implications of representing change for theories of event cognition and event representation. Our intention is to account for the *content* of event representations – or rather, the minimum content that is required to represent an event as might be experienced by an onlooker or conveyed by that onlooker via language. We introduce a new account of event representation based on those aspects of object representations that convey the distinct states that an object has experienced – minimally reflecting the *before* and *after* of whatever changes the object undergoes as an event unfolds: For an event that can be described as "*the chef chopped the onion*", the initial state is an intact onion, and the resultant state is the chopped onion.

Research on event cognition has most recently tended to focus on how we segment the continuous input stream into discrete events (Event Segmentation Theory; Zacks, Speer, Swallow, Braver, & Reynolds, 1997), how information is maintained/recalled within and across event boundaries (e.g. The Event Horizon Model; e.g. Radvansky, 2012), and how to best characterize *generalized event knowledge* – "semantic" knowledge of how, in general, the world changes as a function of typical protagonist interactions (e.g. McRae, Hare, Elman, & Ferretti, 2005; Metusalem, Kutas, Urbach, Hare, McRae, & Elman, 2012). Accounts of event structure and lexical semantics from within the linguistic tradition (e.g. Dowty, 1979; Rappaport Hovav & Levin, 1998; Vendler, 1957; Warglien et al. 2012) posit that the meanings of verbs include reference to objects' changes of state as entailed by the actions denoted by the verb. In all these cases, there is an assumption (tacit or explicit) that actions, participants (objects), space, time (and possibly causality) are primitives of event representation. Embodiment theories of action and object representation (e.g. Barsalou, 1999) likewise appear founded on the principle that action is a primitive, that objects are represented in terms of the perceptions and *inter*actions they afford (Gibson, 1979; Glenberg, 1997), and that more broadly, language itself is grounded in action (e.g. Glenberg & Kaschak, 2002). We contend, however, that action representations are not in fact *primitives* of event representation, but rather are *emergent representations* abstracted across the distinct changes in state that objects undergo in the company of other (changing) objects.

We shall propose the primacy of object-state representations on the basis of a series of studies, from eye-tracking to fMRI, which suggest that when an event is described to have occurred such that an object has changed state, the representations of these distinct states are *simultaneously* activated when that object is subsequently referred to again (cf. homophony, in which a single lexical label can have multiple meanings). Thus, on hearing that the woman chopped the onion, and then hearing "*And then she weighed the onion*", both the initial and resultant states are activated at "*the onion*" (e.g. Solomon, Hindy, Altmann, & Thompson-Schill, 2015). We shall argue for the *action-as-abstraction* approach to event structure on grounds of parsimony: if, as the evidence suggests (and theory requires), the cognitive system represents

[1] Fortuitously, the word *amoeba* is derived from the Ancient Greek for *change*.

the spatiotemporal properties of object states, and does so in the context of other objects' spatiotemporally defined states, what more is required, representationally speaking, to encode action? (We shall confront head-on the challenge of explaining how we reconcile action-as-abstraction with the existence of brain mechanisms that appear specialized for planning and executing motoric action, but space precludes discussion here).

Our account of event representation, that multiple object-states represent events, requires not just the distinction between object types (onions in general) and object tokens (a specific onion in the possession of the woman), but also between tokens and token-states (i.e. the states through which individual tokens pass; the onion in its different, intact and chopped, states). This type/token/token-state distinction places language (and event) comprehension at the interface between semantic memory (e.g. knowledge about onions in general) and episodic memory (e.g. knowledge about a specific onion in a specific state at a particular moment in event-time). We shall describe new evidence that the brain mechanisms implicated in episodic memory (specifically hippocampus; see e.g. Burgess, Maguire, & O'Keefe, 2002, for review) do indeed play a central role in language comprehension; instantiating the tokens and token-states described by the language.

An integral part of event understanding is knowing that an event has occurred, and that the world was in one state but now is in another. The representation of such change, grounded in the interface between semantic and episodic memory, is a fundamental ingredient of language and event comprehension, and indeed, of *cognition* more generally. A fuller understanding of event representation will necessarily require that we cross that semantic/episodic interface.

References

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and brain sciences*, *22*(04), 577-660.

Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, *35*(4), 625-641.

Dowty, D. (1979). Word meaning and Montague grammar. Dordrecht, The Netherlands: Kluwer Academic.

Gibson, J. J. (1979) The ecological approach to visual perception. Houghton- Mifflin.

Glenberg, A. M. (1997). What memory is for: Creating meaning in the service of action. *Behavioral and brain sciences*, *20*(01), 41-50.

Glenberg, A. M., & Kaschak, M. P. (2002). Grounding language in action. *Psychonomic bulletin & review*, *9*(3), 558-565.

McRae, K., Hare, M., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory & Cognition*, *33*(7), 1174-1184.

Metusalem, R., Kutas, M., Urbach, T. P., Hare, M., McRae, K., & Elman, J. L. (2012). Generalized event knowledge activation during online sentence comprehension. *Journal of memory and language*, *66*(4), 545-567.

Radvansky, G. A. (2012). Across the event horizon. *Current Directions in Psychological Science*, *21*(4), 269-272.

Rappaport Hovav, M., & Levin, B. (1998). Building verb meanings. In M. Butt & W. Geuder (Eds.) *The projection of arguments: Lexical and compositional factors*. CSLI Publications.

Solomon, S. H., Hindy, N. C., Altmann, G. T., & Thompson-Schill, S. L. (2015). Competition between mutually exclusive object states in event comprehension. *Journal of Cognitive Neuroscience*, 27, 2324-2338.

Vendler, Z. (1957). Verbs and time. *Philosophical Review*, 56, 143–160.

Warglien, M., Gärdenfors, P., & Westera, M. (2012). Event structure, conceptual spaces and the semantics of verbs. *Theoretical Linguistics*, 38(3-4), 159-193.

Westendorf, C., Negrete, J., Bae, A. J., Sandmann, R., Bodenschatz, E., & Beta, C. (2013). Actin cytoskeleton of chemotactic amoebae operates close to the onset of oscillations. *Proceedings of the National Academy of Sciences*, 110(10), 3853-3858.

Zacks, J. M., Speer, N. K., Swallow, K. M., Braver, T. S., & Reynolds, J. R. (2007). Event perception: a mind-brain perspective. *Psychological bulletin*, *133*(2), 273.

The constructional evolution of grammatical functions and the thematic representation of events

In this paper I explore the possibility of explaining the existence of grammatical functions (GFs) in languages as a consequence of the generalization by learners of correspondences between syntactic form and the representation of the thematic features of events and relations.

The GFs Subject and Object are taken for granted as having descriptive and explanatory significance in the vast majority of grammatical analyses. In some theoretical approaches, such as Relational Grammar, Simpler Syntax, HPSG and LFG they are taken as primitives of syntactic representation. Thus, they are assumed to be universals. In Mainstream Generative Grammar, descended from Chomsky (1965), Subject and Object are simply names given to particular properties of syntactic configurations, but these configurations are assumed to be universal. Even functionalist approaches such as Givón (1997) take the distinction as given even where it is not robustly marked by morphology or word order.

At the same time, there are significant exceptions. For example, Schachter (1976) argued that the grammars of Tagalog and other Phillipine languages are not organized around GFs, but around the marking of Topic and θ-roles. Mithun (1991) shows that there are 'active/agentive' languages that manage the correspondence between morphosyntactically marked grammatical arguments and θ-roles without appeal to GFs, while Mithun (2008) goes a bit further and argues that Central Pomo lacks a Subject GF, although it does have a construction that can be viewed as 'passivization'. There are many other examples in the literature.

The idea that a language can be fully functional in expressing θ-roles and distinguishing arguments without GFs raises the question that I explore in this paper: Where do GFs come from, if they are not universal? And if they are not universal, what are they, and why are they so ubiquitous in the languages of the world?

I hypothesize that the GFs are essentially grammatical categories that arise from the generalization of formal devices for marking thematic properties of arguments. This idea is by no means a new one. For example Shibatani (1991, 103) writes that "our view on the subject is that: (a) it is a syntactic category resulting from the generalization of an agent over other semantic roles, (b) languages vary as to how far this generalization has taken place; i.e. the grammatical status of subject differs from one language to another, and therefore (c) the subject is not necessarily a universal category."

My goal here is to explore the mechanisms that yield the 'generalization' envisioned by Shibatani. The approach thus has three components. First, we must specify as precisely as possible how events and other kinds of relationships between entities are represented conceptually. Focusing (if only for practical reasons) on Subject, this means that we have to say in more precise terms what it means for an argument to have a given thematic role, e.g. AGENT. Fortunately there is a reasonably solid foundation to build on in this regard, in the form of Dowty (1991), the work that he built on, such as Hopper and Thompson (1980), and the work that has responded to his proposals, such as Primus (2002) and Ackerman and Moore (2001), among many others.

Second, we must give a formal characterization of how a grammatical category is formed through generalization. Again, fortunately, there is a long and rich literature of grammaticalization that we can build on. The key idea, which is found through the literature, often implicitly and sometimes explicitly, is that grammatical categories are formed through changes in the conditions that define grammatical constructions (Traugott and Trousdale 2013). I propose a formal characterization of 'construction', a definition of a grammar in constructional terms, and a scenario in which a learner arrives at a grammar by hypothesizing constructions based on input consisting of individual constructs that are licensed by the constructions of the grammar. Generalization (as well as complexification) is a consequence of the imprecision of the learner's hypotheses. I rely heavily on the work of Alishahi and Stevenson (2010), who offer a computational simulation of Bayesian learning of constructions. I suggest that one consequence of imprecise learning is the emergence of differential argument marking (Seržant and Witzlack-Makarevich (submitted, 2015)).

The third component concerns the question of why change actually occurs. Why isn't the imprecision of the learner corrected through subsequent intense exposure to the correct correspondences? I suggest that errors by learners can become established changes if three conditions are satisfied: (i) the errors are not so salient that they demand correction, (ii) the social network in which learners interact has a configuration that allows for errors to spread to a coherent subpopulation, and (iii) the change is computationally advantageous.

References

Ackerman, Farrell, and John Moore. 2001. *Proto-properties and grammatical encoding*. CSLI Publ.

Alishahi, Afra, and Suzanne Stevenson. 2010. A computational model of learning semantic roles from child-directed language. *Language and Cognitive Processes* 25:50–93.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67:547–619.

Givón, Talmy. 1997. Grammatical relations: An introduction. In *Grammatical relations: A functionalist perspective*, volume 35, 1–84. Amsterdam / Philadelphia: John Benjamins Publishing Co.

Hopper, Paul J., and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56:251–299.

Mithun, Marianne. 1991. Active/agentive case marking and its motivations. *Language* 67:510–546.

Mithun, Marianne. 2008. Does passivization require a subject category. *Case and grammatical relations: Studies in honor of Bernard Comrie* 81:211.

Primus, Beatrice. 2002. Proto-roles and case selection in optimality theory. *Arbeiten des SFB* 282:1–39.

Schachter, Paul. 1976. Subject in Philipine languages. In *Subject and Topic*, ed. Charles Li. New York: Academic Press.

Seržant, Ilja A, and Alena Witzlack-Makarevich. submitted, 2015. Differential argument marking: an introduction. In *The diachronic typology of differential argument marking*, ed. Ilja A. Seržant, Alena Witzlack-Makarevich, and K. Mann. Language Science Press.

Shibatani, Masayoshi. 1991. Grammaticization of topic into subject. In *Approaches to grammaticalization*, ed. Elizabeth Closs Traugott and Bernd Heine, volume 2, 93–133. Amsterdam / Philadelphia: John Benjamins Publishing Company.

Traugott, Elizabeth Closs, and Graeme Trousdale. 2013. The framework. In *Constructionalization and constructional change*, ed. Elizabeth Closs Traugott and Graeme Trousdale. Oxford: Oxford University Press.

THE CONSTRUCTIONAL EVOLUTION OF GRAMMATICAL FUNCTIONS AND THE THEMATIC REPRESENTATION OF EVENTS

Peter W. Culicover (The Ohio State University), Afra Alishahi (Tilburg University),
Elena Vaiksnoraite (The Ohio State University)
culicover.1@osu.edu

In this programmatic paper we explore the possibility of explaining the existence of grammatical functions (GFs) in languages as a consequence of the generalization by learners of correspondences between syntactic form and the representation of the thematic features of events and relations.

The GFs Subject and Object are taken for granted as having descriptive and explanatory significance in the vast majority of grammatical analyses. In some theoretical approaches, such as Relational Grammar, Simpler Syntax, HPSG and LFG, they are taken as primitives of syntactic representation. Thus, they are assumed to be universals. In Chomsky (1965), Subject and Object are names given to universal properties of syntactic configurations.

At the same time, there are significant exceptions. For example, Schachter (1976) argued that the grammars of Tagalog and other Phillipine languages are not organized around GFs, but around the marking of Topic and θ-roles. Mithun (1991) shows that there are 'active/agentive' languages that manage the correspondence between morphosyntactically marked grammatical arguments and θ-roles without appeal to GFs, while Mithun (2008) goes a bit further and argues that Central Pomo lacks a Subject GF, although it does have a construction that can be viewed as ?passivization?.. Many languages show 'differential argument marking', which suggests that what we think of as GFs are in fact various degrees of generalization of the correspondence between thematic structure and grammatical form.

The idea that a language can be fully functional in expressing θ-roles and distinguishing arguments without fully general GFs or without GFs at all raises the question that we explore in this paper: Where do GFs come from, if they are not universal? And if they are not universal, what are they, and why are they so ubiquitous in the languages of the world?

The idea that GFs are essentially grammatical categories that arise from the generalization of formal devices for marking thematic properties of arguments is not new. Shibatani (1991:103) writes that "our view on the subject is that: (a) it is a syntactic category resulting from the generalization of an agent over other semantic roles, (b) languages vary as to how far this generalization has taken place; i.e. the grammatical status of subject differs from one language to another, and therefore (c) the subject is not necessarily a universal category." And Keenan (1976) argues that an NP is a subject to the extent that it possesses a subset of properties on a list of semantic, thematic and syntactic properties.

Our goal is to explore the mechanisms that yield the sorts of 'generalization' envisioned by Shibatani. The approach thus has three components. First, we must specify how events and other kinds of relationships between entities are represented conceptually. Focusing here on subject, this means that we have to say in more precise terms what it means for an argument to have

agentive properties. Fortunately there is a reasonably solid foundation to build on in this regard, in the form of Keenan (1976), Dowty (1991) and others, and extensive typological description of differential marking, which we illustrate using data from Agul (Ganenkov et al. 2009).

Second, we must explain how a GF category is formed through generalization. Again, fortunately, there is a long and rich literature of grammaticalization that we can build on. The key idea, which is found throughout the literature, is that grammatical categories are formed through changes in the licensing conditions that define grammatical constructions (e.g. Traugott and Trousdale 2013). We assume a formal characterization of CONSTRUCTION, a definition of a grammar as a set of constructions, and a scenario in which a learner arrives at a grammar by hypothesizing constructions based on input consisting of individual constructs. Generalization (as well as complexification) is a consequence of the imprecision of the learner's hypotheses. We assume Alishahi and Stevenson's (2010) computational simulation of Bayesian learning of constructions to model construction acquisition.

The third component concerns the question of why change actually occurs. Why isn't the imprecision of the learner corrected through subsequent intense exposure to the correct correspondences? We suggest that errors by learners can become established changes if: (i) the errors are not so salient that they demand immediate correction, (ii) the social network in which learners interact allows for errors to propagate in a subpopulation, and (iii) the change is computationally advantageous (Culicover 2013). We propose to adapt the Alishahi and Stevenson (2010) simulation to model change in the network, and hence differential marking, as a consequence of propagation of learner error in the network.

References

Alishahi, Afra, and Suzanne Stevenson. 2010. A computational model of learning semantic roles from child-directed language. *Language and Cognitive Processes* 25:50–93.

Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Culicover, Peter W. 2013. *Grammar and complexity: Language at the intersection of competence and performance*. Oxford: Oxford University Press.

Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67:547–619.

Ganenkov, Dmitry, Timur Maisak, and Solmaz Merdanova. 2009. Non-canonical agent marking in agul. In *Differential subject marking*, ed. Helen de Hoop and Peter de Swart, 173–198. Berlin: Springer.

Keenan, Edward. 1976. Towards a universal definition of "Subject of". In *Subject and Topic*, ed. Charles Li, 303–33. New York: Academic Press.

Mithun, Marianne. 1991. Active/agentive case marking and its motivations. *Language* 67:510–546.

Mithun, Marianne. 2008. Does passivization require a subject category. In *Case and grammatical relations: Studies in honor of Bernard Comrie*, ed. Greville G Corbett and Michael Noonan, 211–40. Amsterdam / Philadelphia: John Benjamins Publishing.

Schachter, Paul. 1976. Subject in Philippine languages. In *Subject and Topic*, ed. Charles Li, 491–518. New York: Academic Press.

Shibatani, Masayoshi. 1991. Grammaticization of topic into subject. In *Approaches to grammaticalization*, ed. Elizabeth Closs Traugott and Bernd Heine, volume 2, 93–133. Amsterdam / Philadelphia: John Benjamins Publishing Company.

Traugott, Elizabeth Closs, and Graeme Trousdale. 2013. The framework. In *Constructionalization and constructional change*, ed. Elizabeth Closs Traugott and Graeme Trousdale, 1–44. Oxford: Oxford University Press.

The incremental processing of accomplishment predicates

Accomplishment predicates like "peeling an apple" denote events that have a duration and a culminating point (Vendler, 1967). When marked with a past perfective form, as in the Italian sentence (1) Valeria $ha_{AUX\text{-}PF}$ $sbucciato_{PAST\text{-}PF}$ la mela [lit. Valeria has peeled the apple], they are considered to be true only if the telos has been reached (i.e., when the apple has been completely peeled in case of (1)). Previous experimental studies show that verb aspect generally constrains situation model construction (Madden & Ferretti 2009; also Frazier et al. 2006) and perfectives yield a representation of a completed event (Ferretti et al. 2009). Nonetheless, telicity inference is shown to be costly: participants do *not* make rapid use of verb/verb+obj information (Proctor et al. 2004); also, comprehenders do not immediately commit fully to the telicity of events (Pickering et al. 2006; Piñango et al. 1999; Todorova et al. 2000). However, these results might depend on the methods employed in those studies, mostly based on behavioural and reading data that might be not sensitive enough to capture this inference. We investigated the online processing of accomplishment verbs in the perfective form in Italian by means of the Visual World eye-tracking paradigm (Tanenhaus et al. 1995), in which participants process sentences in front of a visual context (Altmann and Kamide, 1999). Our aim was to test if and to what extent the telicity inference is computed and/or exploited incrementally upon encountering a telic predicate in the perfective form.

The experiment.

Participants were 29 adult Italian native speakers. They were instructed to find the target between two pictures shown on the screen, following oral instructions in which a lead-in segment like "Dimmi dove…." [lit. Tell me where] was followed by sentences like (1).There were a total of 45 experimental trials and 4 practice trials to illustrate the task. Three conditions were created to test different predictions:

- **Perfective** (PF), Fig. 1: sentences like (1) were shown in front of one completed event (an apple that has been completely peeled) and one incomplete event (an orange that has not been fully peeled yet). In this case, objects were different (apple vs. orange) but type of predicate was compatible with both (peeling)

➔ participants should converge on the picture showing the completed event as soon as the telicity inference was exploited incrementally upon processing the perfective aspect on the verb (Aux+Past Participle);

- **Early Condition** (EC), Fig 2. Object was the same in both pictures (*cornetta* 'receiver') but the predicate was compatible only with one event (has lifted the receiver, Fig. 2-left)

➔people should start to converge on the target possibly before encountering the post-verbal noun, because the predicate was compatible with only one pictured event;

- **Late Condition** (LC), Fig 3. Predicates were non-eventive verbs, sharing initial form *ha* 'has' with the other two conditions. Objects were different (ipod vs. charger) but both were semantically compatible with the predicate (both are hand-held)

➔ people must wait until the post-verbal noun has been processed (or partly processed) because the objects were different and the predicate was semantically compatible with both.

Results.
Fig. 4 shows the timecourse of eye-movements toward the Target picture in the time-window that is settled to start from Aux-onset (time 0) in the three experimental conditions. What it is evident is that in both PF and EC - but not in LC - the target gets significantly more looks than the competitor, suggesting that the preference to look at the target in PF starts before the full Noun is processed. A series of analyses of variance ANOVA show that in EC participants look significantly more to the target in the initial and middle portion of the noun than in the other conditions, as expected (all ps<.01). Also, they look significantly more to the target (the completed event) in PF than in LC in the second and final part of the noun (p<01). We also tested whether the log ratio mean of looks to target/competitor is significantly different from zero at the beginning of the noun and found that the target is looked at significantly more than the competitor in EC and PF (p<.01), but not in LC (p=.97). This gathered indirect information about anticipatory looks to target, i.e. looks occurring before Noun onset.

Conclusions.
We present the first study investigating the online processing of telicity using the visual world paradigm. We found that, when processing an accomplishment predicate marked at the perfective form, comprehenders converged on target (the completed event) before the noun was fully spelt out. Specifically, in the critical condition (PF) we observed a significant preference for the target (over the competitor) already in the very early part of the post-verbal noun, while this preference emerged later in LC. Contra previous studies, our results suggest a relatively early and easy integration of the telicity inference (triggered by perfective aspect on verb) during online incremental processing. Future experiments are planned to compare perfective and imperfective aspect with a similar paradigm.

Fig. 1 Perfective (PF)
ha sbucciato la mela
[has peeled the apple]

Fig. 2 Early-Control (EC)
ha sollevato la cornetta
[has lifted the receiver]

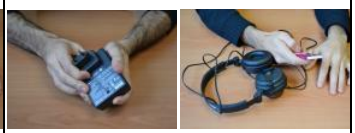
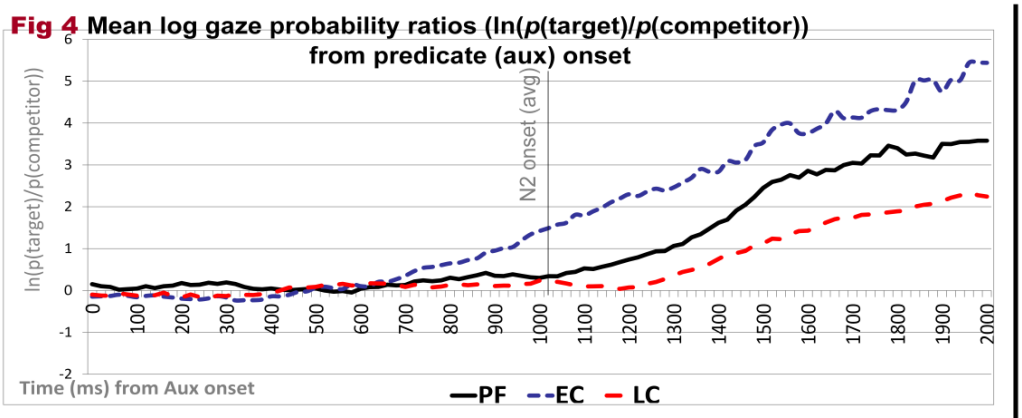


Fig. 3 Late-Control (LC)
ha in mano un ipod
[…has in hand an ipod]



Grammatical constraints on event packaging and potential effects on the segmentation of the perceptual stream: Motion events in language and cognition under a cross linguistic perspective

Definitions of the category *event* can be found in different disciplines, ranging from linguistics, cognitive science, psychology to physics. They all have in common that they take events as entities in time (cf. Davidson's definition 'events are objects in time' or in physics as 'a single occurrence that is localized at a single point in space and instant of time'). What we do not get by this definition, however, is a criterion for segmentation, i.e., for setting boundaries. Previous studies on event segmentation in cognition (e.g. Radvansky & Zacks 2014) and linguistics (e.g. Bohnemeyer & Pedersen 2011) show that there is no 'natural' way of forming events. While we know some of the criteria people draw on when forming events, such as *change of state of an entity* or *causal relations between states of affaires*, the challenging questions remain which factors guide a person a) when segmenting the perceptual stream in cognitively manageable packages and b) when selecting components to form a propositional unit, expressed in language, and c) whether and d) how these levels interact. It is the latter question which we will address in our paper: We are interested in the role of linguistic structure in processes of segmentation in perception and cognition. The conceptual domain of motion events provides an interesting field of study in this respect. It is well-known that languages package information on motion events differently (cf. Talmy 2000). Cross linguistic research has shown that different languages do not only offer different options for referring to motion events but that event construal is grammatically constrained leading to differences in what constitutes a *reportable event* in a language (the authors). The following examples illustrate two types of constraints: (1) path segmentation; *A child is running across the lane into the house.*vs. *Un enfant traverse le chemin et entre dans la maison.*('A child crosses the lane and enters the house'). Both verbalisations refer to the same limited part of the external world. English speakers, however, can present it as one unit of cognitive processing, while French speakers must form two, if they want to express the same amount of information. (2) Boundary crossing; *A car is driving into the village* vs. *Une voiture roule (sur la route) et entre dans le village*. (A car is driving (on the road) and is entering the village.). Again, we have a situation which can be expressed in one cognitive unit in English, including manner and path of motion, while French speakers have to produce two.

Drawing on a large data base on the description of motion events across nine languages (the authors, 2006-2015), elicited by using short real world video clips as stimuli, we can identify specific types of scenes which are almost uniformly segmented differently by speakers of different languages, patterns which we attribute to the different linguistic systems. The question we will approach in the current study is the following: Are the differences in event representation at the verbal level (type 1 and 2 above) reflected in visual processing of the input? If so, are there differences between segmentation for verbalization and segmentation in a non-verbal task? The methodology we use is eye tracking during scene description (verbal) and during a button-press (non-verbal) segmentation task. First results reveal that native speakers of French and English show different segmentation patterns which to some extent correlate with specific patterns of visual attention. The results will be discussed in the

context of the general question on principles of event segmentation across cognitive modes.

References

Jürgen Bohnemeyer and Eric Pederson. (2011). Event Representation in Language and Cognition. Cambridge, MA: CUP

Gabriel A. Radvansky and Jeffrey M. Zacks (2014). Event Cognition. Oxford: OUP.

Leonard Talmy (2000). Toward a Cognitive Semantics. Cambridge, MA: MIT Press.

CROSSLINGUISTIC VARIATION IN EVENT CONCEPTUALIZATION: EVIDENCE FROM THE CAUSATIVE ALTERNATION

Björn Lundquist (University of Tromsø), Gillian Ramchand, Mai Tungseth (University of Tromsø), Martin Corley and Antonella Sorace (University of Edinburgh)
bjorn.lundquist@uit.no

What is the semantic relation between causative (C: *John melted the butter*) and anticausative (AC: *the butter melted*) descriptions? And is the semantic relation between C and AC the same across different languages? Although many languages overtly mark the alternation in one way or another (either marking C, AC or both, see Haspelmath 1993), the effect that such morphological markings might have on the semantics of causation is not well understood. There are two broadly competing hypotheses for expressing the semantic relationship between the C verb and its AC alternant:
The Causational Hypothesis: The AC verb is the C minus either a causer, or an entire cause predicate (Grimshaw 1982, Levin and Rappaport 1995, Reinhart and Siloni 2005).
The Reflexive Hypothesis: The AC (or at least overtly marked ACs) are literally reflexive versions of Cs (Chierchia 2004, Koontz-Garboden 2009), where the sole argument of AC is associated both with a theme/patient role and a cause role.

In addition, researchers disagree on whether there should be a language-general answer to the question (the Universalist stance, e.g., Horvath & Siloni, 2011) or whether languages can differ in how the semantic relation between A and AC is instantiated (the Language-specific stance), with a possible correlation to morphological marking.

The Causational Hypothesis predicts that C always entails AC since AC is semantically a less specified version of C. An entailment test is often used to support this hypothesis: *John melted the butter, #but the butter didn't melt.* By contrast, according to the Reflexive Hypothesis, C doesn't entail AC, because AC in this case contains semantic information not present in C; more specifically, that the sole argument is not only undergoing change, but is also responsible for the initiation of the event. However, despite the lack of entailment relation between A and AC, the two descriptions could felicitously describe the same event or state: the sentences, (a) *the barber on the corner cut my hair yesterday* and (b) *I cut my hair yesterday* could both be used to describe the same event. The felicity of the (b) description is predicted to be influenced by a number of factors, if the reflexive hypothesis is correct:(i) is the external argument in (b) enough involved in the event to qualify as a "cause" and (ii) is the direct causer (here, 'the barber') peripheral enough to leave out? The reflexive hypothesis thus predicts that negating AC while claiming C should not be contradictory (e.g, *the door didn't open – I opened it*). However, it has proved difficult to secure consistent judgements of entailment because of interference from pragmatic factors such as metalinguistic negation (see Schäfer & Vivanco for a recent review).

Experiment. In the present experiment, we use video clips to directly pair event depictions with linguistic representations, thus avoiding confounds such as negation. 42 Norwegian and 44 English informants watched short video sequences, in which causative events were depicted (e.g., a woman rolling a ball across the road). Each sequence was followed by an AC question (e.g., *did the ball roll across the road*), and response (yes or no) was logged. In Norwegian, 7 of the video sequences were based on verbs that had form-identical C and AC forms (e.g., *rulle* 'roll', *smelte* 'melt'). The remaining 7 sequences were based on verbs which require the simple reflexive pronoun *seg* in the AC frame (e.g., *åpne seg* 'open', *dele seg* 'split'). Under the Causational Hypothesis, uniform 'yes' responses are expected. Under the Reflexive Hypothesis, a *Yes* answer is not uniformly expected, but it is important to note that *No* answers across the board are not expected either. Rather, responses should be sensitive to manipulations involving degree of agent-involvement/saliency of theme involvement. Therefore, to further test the Reflexive Hypothesis, two video sequences were made for each verb: one where the agent acted intentionally and was active throughout the event (Agent Focus), and one where the focus was on the theme rather than the agent (for example, where the agent was only involved at the

starting point of the event, or caused the event by accident: Theme Focus).

All videos were also tested with a causative question (e.g. *did the woman roll the ball across the road*) with a separate group of speakers ($n = 20$, giving over 93% 'yes' responses) to make sure that the verbs correctly described the depicted events. Further, the anti-causatives were checked with native speakers and against corpora, to make sure that the verbs in question could in principle be used in an anti-causative frame. In English, the same 28 videos were used, corresponding to 14 verbs which are all unmarked. Each informant in each language saw only one version of each video, and thus saw 14 video sequences in all, balanced between Agent-Focus and Theme-Focus contexts, and (for Norwegian) marked and unmarked anti-causatives. An additional 22 filler videos with questions were also presented, and order of presentation was randomised. *The Causational hypothesis predicts that informants will answer 'yes' to all questions, while the Reflexive hypothesis predicts that 'yes' responses will be variable, depending on degree of agent-involvement.*

Results. English-speaking informants were almost at ceiling, giving 'yes' responses 90.5% of the time (numbers in table 1). The effect of context was very small, but significant, in English (6.8%, $p < .05$), with the effect located to only 2 of the 14 predicates. We conclude that, with possible lexical exceptions, AC descriptions of caused events are licit in English even when the focus is on the Agent, just as Hypothesis 1, but not 2, predicts. The Norwegians' total of 'yes' responses was significantly lower the the English (64.4% of trials $p < .001$). The effect of context was very strong in Norwegian: 20.8% ($p < .001$) more 'yes' responses in the Theme focus condition. Norwegian informants showed an effect of *seg* marking and a marking-by-context interaction ($p < .05$), such that context effects were stronger for verbs which don't take *seg*, but the overall proportion of 'yes' responses were bigger for them. A closer look at the materials suggests that reflexive-marked verbs require the agent to be highly peripheral (accidental cause) for the AC interpretation to be felicitous. This suggests that the Reflexive hypothesis is correct for Norwegian, both for unmarked and for marked ACs.

Our experiment thus shows that a universalist position on the analysis of the causative-inchoative alternation across languages is untenable. We have to conclude that natural languages can make use of either "anti-causative" strategies, *or* a reflexive strategy. We suggest that the difference between the two languages has its origin in the fact that AC verbs in Norwegian often carry the same morphological marking as reflexive verbs, which has led speakers to impose the argument structure of reflexive verbs onto AC verbs quite generally. However, the specific factor of reflexive marking does not itself appear to be criterial, since the non-reflexive alternators in Norwegian also showed a sensitivity to the contextual manipulation not displayed by English. The results suggest that speakers' conceptualization of events may be influenced by the *overall shape of the linguistic system.* They also show that fine-grained semantic differences between translational equivalents are easy to overlook in linguistic descriptions. Direct experimental testing of nonlinguistically presented events is necessary to uncover these differences.

Table 1: Percent 'yes' responses, by context

	Eng. total	Nor. total	Nor. unmarked	Nor. refl. marked
Agent Focus	87%	54%	63%	45%
Theme Focus	93%	75%	90%	59%

CAN INTRANSITIVE CLAUSES NAME 2-PARTICIPANT EVENTS?
A NEW TEST OF PARTICIPANT-TO-ARGUMENT MATCHING IN VERB LEARNING

Laurel Perkins (University of Maryland), Angela Xiaoxue He (Boston University), Alexander Williams (University of Maryland), Rachel Dudley (University of Maryland), Sigríður Björnsdóttir (University of Maryland), Jeffrey Lidz (University of Maryland)
Contact: Laurel Perkins, perkinsl@umd.edu

On one verb-learning hypothesis, children expect the number of participants in an event to match one-to-one the number of arguments of a verb describing that event (Participant-to-Argument Matching, PAM) [1-3]. However, previous studies have found inconclusive evidence for PAM with intransitive sentences [3, 4]. The current study provides an improved test for PAM: we (i) introduce a new method to test the fit between a sentence and a scene, (ii) ensure that presented sentences are parsed as intransitive, and (iii) ensure that presented events are perceived with 2 participants. We preliminarily find that children do not consider an intransitive sentence to be a good fit for a 2-participant event, a result consistent with PAM for intransitives.

In previous preferential looking studies, children who hear a novel transitive verb are more likely to look at an event intended to be viewed with 2 participants rather than 1 [1, 3], but children who hear a novel intransitive verb do not reliably show above-chance looking times to an event intended to be viewed with 1 participant rather than 2 [3, 4]. There are several potential reasons for the inconclusive results with intransitives. Children may not perceive the presented sentence as intransitive [3, 5], or children may not perceive the event with the intended number of participants—an event intended to be viewed as one person pushing another might also be viewed as two people playing [6, 7]. Furthermore, the preferential looking task measures which interpretations of a sentence children prefer, rather than which interpretations they have available, which makes it difficult to differentiate PAM from a weaker strategy. Children might only expect that each Argument Names a Participant (ANP), but not that every participant necessarily corresponds to an argument [8]. This weaker heuristic would still be consistent with the previous results. In order to differentiate PAM from ANP, a new method is necessary.

To this end we introduce a Violations of Expectations task [9]. This task measures how well a scene and a sentence fit together, rather than which of two scenes children prefer. We familiarize children with a 2-participant event with uninformative audio, and then measure how much a transitive or intransitive sentence violates children's expectations after viewing this event. PAM predicts that children will be surprised to hear a 2-participant event labelled by an intransitive sentence, whereas ANP does not. To be confident that children perceive our intransitive sentences as intransitive, we use simple pronominal subjects [3]. To be confident that children perceive our videos under a 2-participant event concept, we present brief events in which a human actor effects a sudden change in an inanimate patient: a woman tearing a piece of paper in half, or knocking over a tower of toys. A norming study with adults supported our assumption that these videos are likely to be perceived with 2 participants, not 1.

The current study tests 19- to 22-month-old children in a 2x2 design, with event (TEAR or KNOCK OVER) as a within-subjects factor and clause type (transitive or intransitive) as a between-subjects factor. Pairing of event and novel verb (*blick* and *gorp*), as well as order, are counterbalanced across participants. During familiarization, children see 4 trials each consisting of a 5.4-second event repeated up to 5 times, accompanied by uninformative audio. If a child looks away for more than 2 seconds, the trial is stopped and the next trial begins. During the test phase, each child sees 2 trials with the same video stimuli, but different audio: each child hears either transitive or intransitive sentences with a novel verb. The dependent variable is looking time during the test phase. See Figure 1 for a diagram of the experimental set-up.

Data collection is currently underway. Figure 2 reports preliminary mean looking times during the last familiarization trial and first test trial collapsed across both events, for a current sample of 17 participants out of 32 in our expected final sample. At this stage we find no difference

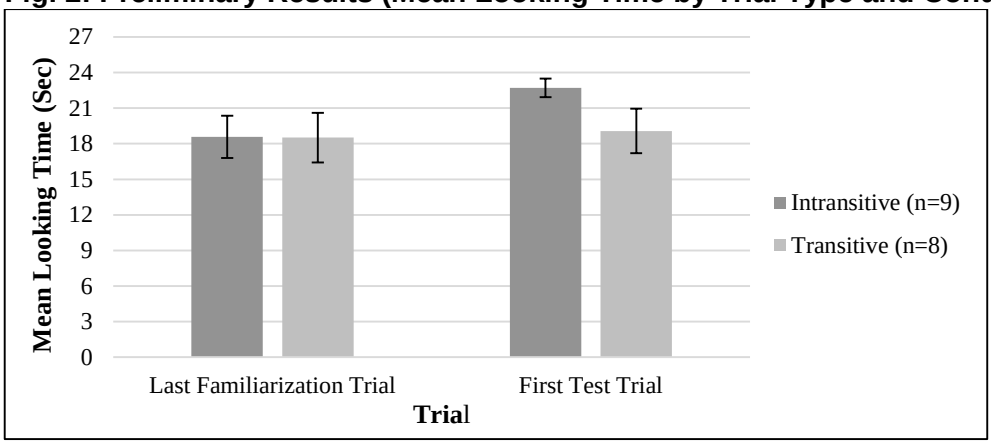
between conditions during familiarization, but during the first test trial we find longer looking time in the intransitive condition compared to the transitive condition, trending towards a significant difference (p = 0.10, t = 1.80). We take this as a preliminary indication that children are initially surprised to hear intransitive sentences label 2-participant events.

These preliminary results are consistent with PAM and inconsistent with the weaker ANP: children expect the 2-participant events in our stimuli to be described with a transitive, and find an intransitive to be a poor fit. If this trend continues in the expected direction, it will provide new evidence that children use a learning strategy stronger than ANP, at least for 2-participant events of the sort we tested here. Further investigation is needed to determine whether this strategy is PAM. Because PAM says that participants must match arguments for any number, we aim in a concurrent project to test this prediction by pitting 3-participant events against transitive sentences. The task introduced in the current study may provide an improved method for future research to differentiate PAM from weaker possible strategies in verb learning.

Fig. 1: Experimental Set-Up (1 of 2 Videos)

	Video	***Audio***	
		Intransitive Condition	Transitive Condition
Familiarization (4 trials)	girl tears paper	*Hey, wow! Do you see that?*	
Test (2 trials)	girl tears paper	*She's gonna blick! She just blicked!*	*She's gonna blick it! She just blicked it!*

Fig. 2: Preliminary Results (Mean Looking Time by Trial Type and Condition)



Selected References

[1] Naigles 1990. Children use syntax to learn verb meanings. *J Child Lang* 17. [2] Fisher et al. 2010. Syntactic bootstrapping. *WIRES Cog Sci* 1. [3] Yuan, Fisher, & Snedeker 2012. Counting the nouns. *Child Dev* 83. [4] Noble, Rowland, & Pine 2011. Comprehension of argument structure and semantic roles. *Cog Sci* 35. [5] Gertner & Fisher 2012. Predicted errors in children's early sentence comprehension. *Cognition* 124. [6] Brandone et al. 2006. One-for-one and two-for-two. *Proc BUCLD 30*. [7] Pozzan, Gleitman, & Trueswell 2015. Semantic Ambiguity and Syntactic Bootstrapping. *Lang Learn Dev* 00. [8] Williams 2015. *Arguments in Syntax and Semantics*. [9] Baillargeon, Spelke, & Wasserman 1985. Object permanence in five-month-old infants. *Cognition* 20.

WHEN THE FOOD ARRIVES BEFORE THE MENU:
MODELING EVENT-DRIVEN SURPRISAL IN LANGUAGE COMPREHENSION

Noortje Venhuizen, Harm Brouwer and Matthew Crocker (Saarland University)
{noortjev,brouwer,crocker}@coli.uni-saarland.de

Online language comprehension is strongly guided by our knowledge about the world, including our knowledge about structured events [e.g., 1, 2]. One way to think about this event knowledge is in terms of the typical co-occurrence and ordering of basic everyday events (cf. the idea of *scripts* [3] or *schemata* [4]). For instance, our knowledge about *going to a restaurant* typically consists of a sequence of events, including 'getting the menu', 'ordering dinner', 'eating' and 'paying'. During linguistic processing, this kind of event knowledge is readily used by comprehenders, for instance to predict upcoming words, as illustrated in the following example: *"Tina asked for the menu and … [ordered/ ate/ left]"*. In this example, the preference for 'ordered' as opposed to 'ate' and 'left' critically depends on our knowledge about the temporal ordering of events in a restaurant setting, and cannot be deduced from the linguistic context alone. However, due to the difficulty in representing and incorporating world knowledge in a computational model of linguistic processing, this information is generally disregarded [5]. A case in point is a recent strand of models that estimate word processing difficulty in online language processing through information theoretic metrics such as *surprisal* [6,7] and *entropy reduction* [8]. The core idea of these models is that the difficulty in processing a word w_t can be estimated as a function of its probability given its prior linguistic ($w_1 \ldots w_{t-1}$) and extra-linguistic context (ELC): e.g., $surprisal(w_t) = -\log P(w_t|w_1 \ldots w_{t-1}, \text{ELC})$. The probability of w_t given the linguistic context can be straightforwardly derived from language models, but the contribution of the extra-linguistic context (and associated world knowledge) is more difficult to quantify and therefore generally left out of implemented models.

In order to improve upon this situation, we present a neurocomputational—recurrent artificial neural network—model of language processing that integrates linguistic knowledge and world/event knowledge, and that produces word surprisal estimates that take into account both. Our model constructs a cognitively motivated situation model of the state-of-the-affairs as described by a sentence [9]. Crucially, these situation model representations inherently encode world/event knowledge. We will show that the surprisal estimates that our model produces reflect both linguistic surprisal as well as surprisal that is driven by knowledge about structured events. We will outline how we can employ the model to explore the interaction between these types of knowledge in online language processing.

Situation model representations encoding event structure

Modeling the interaction between linguistic and event knowledge within a recurrent neural network model, requires a representational scheme for meaning that inherently encodes knowledge about event structure. To this end, we employ the Distributed Situation Space (DSS) model: a psychologically motivated, distributed representational scheme for meaning. DSS representations have been successfully employed to model aspects of both story comprehension [10] and world-knowledge driven semantic processing [9]. In the DSS model, an observation of a state of affairs is represented as a vector encoding which events are the case in that state of affairs. A large sample of such observations creates a high-dimensional "situation-state space", which encodes all knowledge about the world by means of event co-occurrences. To render this feasible, the situation-state space encodes a *microworld*: a confined world consisting of a limited set of events and conditions on the co-occurrence of these events.

The DSS model defines a microworld in terms of a finite set of atomic events (e.g., *ask_menu(tina)*). Observations in a microworld are described in terms of these atomic events, which can be combined to form complex events (e.g., *ask_menu(tina)* $\wedge$ *order(tina,dinner)*). Critically, world knowledge dictates that not all combinations of events are possible (hard constraints) or equally likely (probabilistic contraints). A situation-state space, then, is a large sample of independent, (mostly) complex observations. As such, there is no temporal

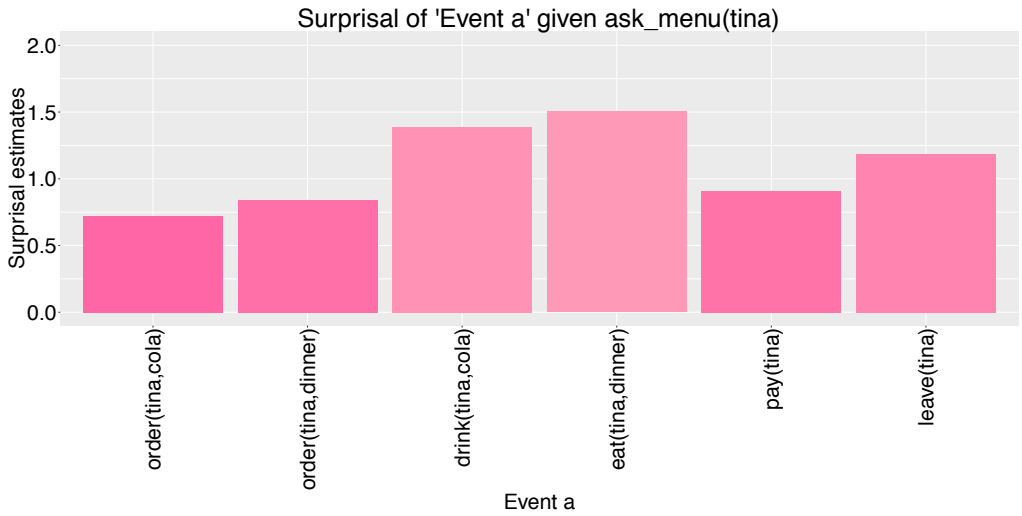


Figure 1: Surprisal estimates from the DSS model for a set of events given 'ask_menu(tina)'.

dependence between observations. In order to incorporate temporal ordering between events within DSS representations, we restrict each observation to contain a temporally continuous sequence of events; i.e., any observation that contains the events *get_menu(tina) and eat(tina, dinner)* must also contain the event *order(tina,dinner)*. Importantly, the "situation-state space" inherently carries information about the likelihood of events in the world relative to other events [see 10]. This allows for estimating the world knowledge-driven surprisal of an event given any other (possibly complex) event. Figure 1 illustrates the world knowledge-driven surprisal estimates for a set of events in our microworld, given the event *ask_menu(tina)*. This figure shows that the temporally closer events (*order(tina,cola)* and *order(tina,dinner)*) are more expected in this context (i.e., obtain lower surprisal) than later events such as 'drinking', 'eating', 'paying', and 'leaving'. Moreover, the relatively high surprisal for *eat(tina,dinner)* and *drink(tina,cola)* as compared to *pay(tina)* and *leave(tina)* illustrates the probabilistic nature of the microworld; in a restaurant, 'eating' and 'drinking' is more optional than 'paying' and 'leaving'.

Modeling the online interaction between linguistic and event knowledge

In order to obtain an online measure of surprisal, which is driven by both the linguistic context and event knowledge encoded in the DSS, we employ a Simple Recurrent Neural network (SRN) of the Elman type [11]. The model is trained to take sequences of word representations as input (*linguistic experience)*, and to produce a situation model of the state-of-the-affairs described by a sentence—a DSS representation encoding *world knowledge*—on a word-by-word basis. Hence, the model naturally combines expectations due to linguistic experience with world/event knowledge. The surprisal of a word w_t, then, can be estimated directly from the situation model sm_t constructed after processing w_t, and the previous situation model sm_{t-1} constructed based on $w_{1...}w_{t-1}$; $surprisal(w_t) = -\log P(sm_t | sm_{t-1})$. These estimates allow us to model phenomena in which event knowledge has been shown to influence expectations above and beyond purely linguistic surprisal. These include discourse level event representations [2], the influence of event knowledge on verb complement expectations [12], the processing of coercion phenomena [13], and the effects of temporal violations in event sequences [14].

References. [1] Zacks & Tversky (2001), *Psychol. Bull.*; [2] Metusalem et al. (2012), *JML*; [3] Schank & Abelson (1977); [4] Rumelhart (1980), *In: Spiro et al. (eds)*; [5] Frank et al. (2008), *Discourse Process.*; [6] Hale (2001), *NAACL*; [7] Levy (2008), *Cognition*; [8] Hale (2006), *Cognitive Sci.*; [9] Frank et al. (2009), *Cognition*; [10] Frank et al. (2003), *Cognitive Sci.*; [11] Elman (1990), *Cognitive Sci.*; [12] Bicknell et al. (2010), *JML*; [13] Kuperberg et al. (2010), *JOCN*; [14] Raisig et al. (2010), *Int. J. Psychophysiol.*

When the food arrives before the menu
Modeling Event-driven Surprisal in Language Comprehension

Online language comprehension is strongly guided by our knowledge about the world, including our knowledge about structured events [e.g., 1, 2]. One way to think about this event knowledge is in terms of the typical ordering of basic everyday events (cf. the idea of *scripts* [3] or *schemata* [4]). For instance, our knowledge about *going to a restaurant* typically consists of a sequence of events, including 'getting the menu', 'ordering food', 'eating' and 'paying'. During linguistic processing, this kind of event knowledge is readily used by interpreters, for instance to predict upcoming words, as illustrated in the following example: *"Rita entered a restaurant and asked for … [the menu/ the bill]"*. In this example, the preference for 'the menu' as opposed to 'the bill' critically depends on our knowledge about the temporal ordering of events in a restaurant setting, and cannot be deduced from the linguistic context alone. However, due to the difficulty in representing and incorporating world knowledge in a computational model of linguistic processing, this information is generally disregarded [5]. A case in point is a recent strand of models that estimate word processing difficulty in online language processing through information theoretic metrics such as *surprisal* [6,7] and *entropy reduction* [8]. The core idea of these models is that the difficulty in processing a word w_t can be estimated as a function of its probability given its prior linguistic ($w_1 \ldots w_{t-1}$) and extra-linguistic context (ELC): e.g., $surprisal(w_t) = -\log P(w_t | w_1 \ldots w_{t-1}, \text{ELC})$. The probability of w_t given the linguistic context can be straightforwardly derived from language models, but its probability given the extra-linguistic context is generally left out of implemented models.

In order to improve upon this situation, we present a neurocomputational—recurrent artificial neural network—model of language processing that integrates linguistic knowledge and world/event knowledge, and that produces word surprisal estimates that take into account both. Our model constructs a cognitively motivated situation model of the state-of-the-affairs as described by a sentence [9]. Critically, these situation model representations inherently encode world/event knowledge. We will show that the surprisal estimates that our model produces reflect both linguistic surprisal as well as surprisal that is driven by knowledge about structured events. We will outline how we can employ the model to explore the interaction between these types of knowledge in online language processing.

Situation model representations encoding event structure

Modeling the interaction between linguistic and event knowledge within a recurrent neural network model, requires a representational scheme for meaning that inherently encodes knowledge about event structure. To this end, we employ the Distributed Situation Space (DSS) model: a psychologically motivated, distributed representational scheme for meaning. DSS representations have been successfully employed to model aspects of both story comprehension [10] and world-knowledge driven semantic processing [9]. In the DSS model, an observation of a state of affairs is represented as a vector encoding which events are the case in that state of affairs. A large sample of such observations creates a high-dimensional "situation-state space", which encodes all knowledge about the world by means of event co-occurrences. To render this feasible, the situation-state space encodes a *microworld*: a confined world consisting of a limited set of events and conditions on the co-occurrence of these events.

The DSS model defines a microworld in terms of a finite set of atomic events (e.g., *enter(rita,restaurant)*). Observations in a microworld are described in terms of these atomic events, which can be combined to form complex events (e.g.,*enter(rita,restaurant)* $\wedge$ *leave(rita,restaurant)*). Critically, world knowledge dictates that not all combinations of events are possible (hard constraints) or equally likely (probabilistic contraints). A situation-

Figure 1: Surprisal estimates from the DSS model for a set of events given 'enter(rita,restaurant)'.

state space, then, is a large sample of independent, (mostly) complex observations. As such, there is no temporal dependence between observations. In order to incorporate temporal ordering between events within DSS representations, then, we assume that an observation must contain the temporally preceding events for any given event; i.e., an observation that contains the event *get_menu(rita,restaurant)* must also contain the event *enter(rita,restaurant)*. Critically, the "situation-state space" inherently carries information about the likelihood of events in the world relative to other events [see 10]. This allows for estimating the surprisal of an event given any other (possibly complex) event. Figure 1 illustrates the surprisal estimates for a set of events in our microworld, given the event '*enter(rita,restaurant)*'. This figure shows that the temporally closer event '*get_menu(rita,restaurant)*' is more expected than '*get_bill(rita,restaurant)*' in this context.

Modeling the online interaction between linguistic and event knowledge

In order to obtain an online measure of surprisal, which is driven by both the linguistic context and event knowledge encoded in the DSS, we employ a Simple Recurrent Neural network (SRN) of the Elman type [11]. The model is trained to take sequences of word representations as input (*linguistic knowledge)*, and to produce a situation situation model of the state-of-the-affairs—a DSS representation—described by a sentence on a word-by-word basis (*event knowledge*). Hence, the model naturally combines linguistic and situational knowledge, which allows for estimating the surprisal of a word w_t directly from the situation model sm_t that is constructed after processing w_t, and the previous situation model $sm_{1...t-1}$: $surprisal(w_t) = -\log P(sm_t | sm_{1...t-1})$. Using these estimates, we can model the effect of different phenomena in which event knowledge has been shown to influence expectations above and beyond purely linguistic surprisal. These phenomena include discourse level event representations [2], the influence of event knowledge on verb complement expectations [12], the processing of coercion phenomena [13], and the effects of temporal violations in event sequences [14].

References. [1] Zacks & Tversky (2001), *Psychol. Bull.*; [2] Metusalem et al. (2012), *JML*; [3] Schank & Abelson (1977); [4] Rumelhart (1980), *In: Spiro et al. (eds)*; [5] Frank et al. (2008), *Discourse Process.*; [6] Hale (2001), *NAACL*; [7] Levy (2008), *Cognition*; [8] Hale (2006), *Cognitive Sci.*; [9] Frank et al. (2009), *Cognition*; [10] Frank et al. (2003), *Cognitive Sci.*; [11] Elman (1990), *Cognitive Sci.*; [12] Bicknell et al. (2010), *JML*; [13] Kuperberg et al. (2010), *JOCN*; [14] Raisig et al. (2010), *Int. J. Psychophysiol.*

EVENT ROLE IDENTIFICATION IN LANGUAGE AND COGNITION

Ercenur Ünal (University of Delaware), John Trueswell (University of Pennsylvania),
Anna Papafragou (University of Delaware)
eunal@udel.edu

Much of the literature assumes that the expression of events in language is based on the salience of event components at the perceptual/conceptual level.[1-4] However, with the exception of a few studies,[5-7] little empirical work has tested this assumption directly. Furthermore, much of the evidence on the relative prominence of event components comes from English. Here, we ask whether the asymmetries in event role prominence generalize to other languages drawing on evidence from learners of two languages (English and Turkish) which vary in how often and the ways in which they encode event components.

Stimuli were 24 pictures of caused-motion events in which a person/animal (Agent) used a tool/body part (Instrument) to move an object (Patient) to a destination (Goal) (Fig.1). In English, Agents typically appear as Subject NPs, Patients as Direct Object NPs, Goals as PPs and Instruments as PPs or Vs (e.g., *raking*).[7] Turkish shows more surface variation due to pro-drop and case marking: in addition to English-type possibilities, the Subject can be dropped and Goals and Instruments can appear as case-marked NPs. Two experiments sought to confirm these cross-linguistic differences (Exp.1) and ask whether English and Turkish speakers prioritize event components in the same way in language (Exp.1) and cognition (Exp.2). Animate (Agents) and inanimate entities were treated separately.

Experiment 1 assessed the prominence of event components in language. Twelve 3-year-olds and twelve 4-year-olds from each language group described the action depicted in 14 of the pictures. As expected, Agents appeared as Subject NPs in both languages, but were mentioned less in Turkish (Fig.2). Patients were always encoded as Direct Object NPs. For the remaining components, as expected, descriptions diverged: Turkish learners used case-marked NPs frequently for Goals and exclusively for Instruments. Despite this syntactic variation, both age and language groups mentioned Patients and Goals more than Instruments (even though 4-year-old English speakers additionally mentioned Patients more often than Goals; Fig.2).

Experiment 2 assessed the cognitive salience of event components using a change-blindness paradigm in which two versions of a picture flickered with a gray mask in-between. The color of one event component changed between versions. Of interest was whether change detection accuracy would differ for each component and whether the differences would be modulated by language. Children in both age and language groups had equally low accuracy rates (M=0.48) for Agent-changes, probably because the color change in Agents affected a property (i.e., the clothes) of the component, not the component itself. All children were more accurate in detecting Patient- and Goal-Changes than Instrument-Changes (Fig.3). Four-year-olds were overall more accurate than three-year-olds ($F(1,24)=12.62, p=.002$).

Summarizing, in both language and cognition, Patients and Goals tended to be prioritized over Instruments. Furthermore, Turkish children who mentioned Agents less frequently in language were no less accurate than English children in detecting Agent-changes. The relative salience of event components was similar across young learners of English and Turkish, despite the cross-linguistic differences in the syntactic positions in which entities were encoded. These findings offer cross-linguistic evidence for a homology between linguistic and conceptual event roles.

References
1. Jackendoff, R.S. (1990). *Semantic Structures*. Cambridge, MA: MIT Press.
2. Baker, M.C. (1997). Thematic roles and syntactic structure. In L. Haegeman (Ed.), *Handbook of Generative Syntax* (pp. 73-137). Dordrecht: Kluwer.
3. Levin, B., & Rappaport Hovav M. (2005). *Argument realization*. Cambridge, UK: Cambridge University Press.
4. Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure.* Cambridge, MA: MIT Press.
5. Lakusta, L. & Landau, L. (2012). *Cognitive Science,* 36, 517-544.
6. Papafragou, A. (2010). *Cognitive Science, 34*, 1064-1092.
7. Authors. (2014) BUCLD39.

Figure 1. Example event ("A man hit a ball into a bucket with a golf club")

Figure 2. Proportion of mention of each event component across Age and Language groups (Exp.1)

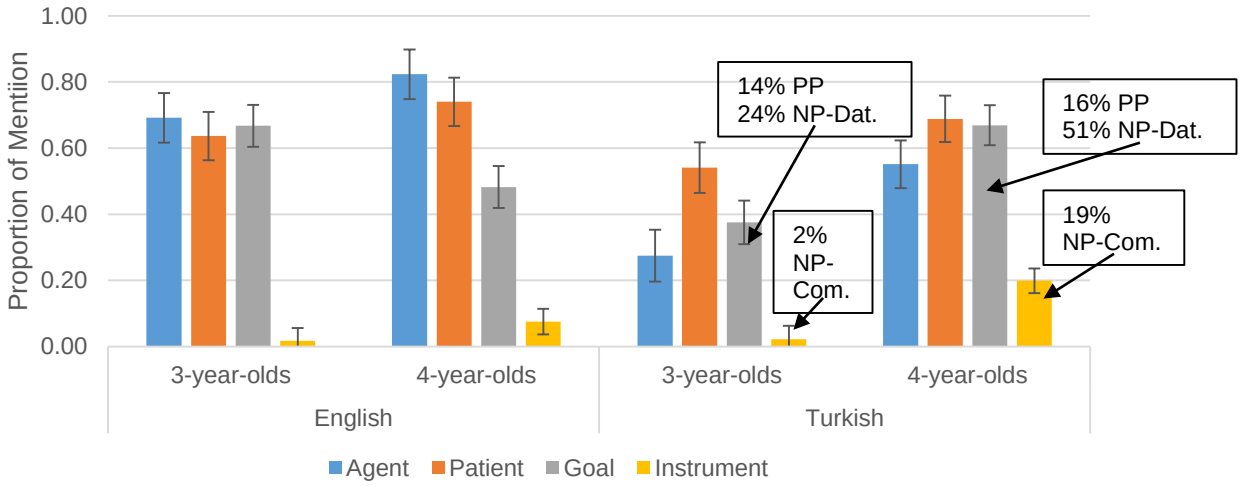
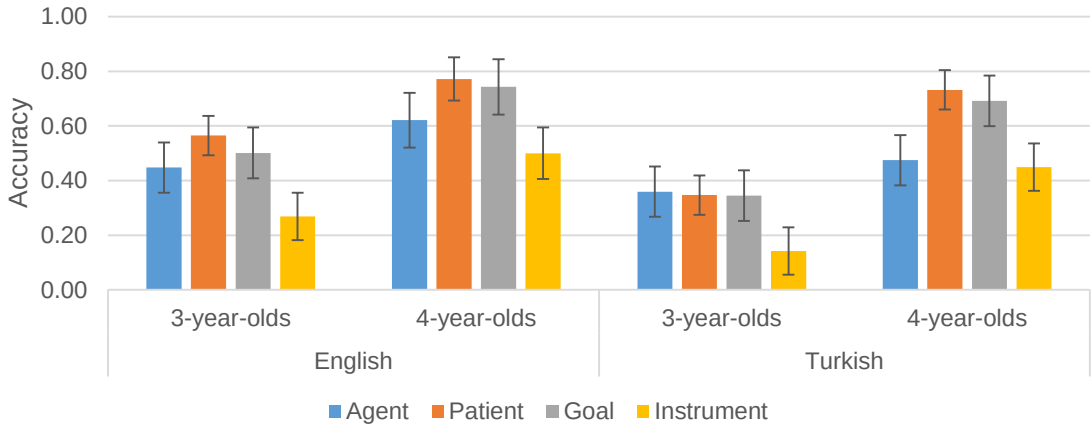


Figure 3. Accuracy in detecting changes in each event component across Age and Language groups (Exp.2)



EVENT PARSING AND VERB LEARNING:
HOW WELL DO CHILDREN LEARN VERBS AMIDST DISTRACTIONS?

Jane Childers (Trinity University), Shelly Gordon and Lupita Mercado (Trinity University)
jchilder@trinity.edu

Verb acquisition requires children to segment dynamic scenes and link different elements to specific verbs. In most verb studies, children only see relevant events. Two studies ask 1) how well children learn verbs while seeing relevant and irrelevant events, and 2) if they can parse events, linking only relevant subevents to new verbs. Structural alignment (SA) theory (e.g., Markman & Gentner, 1997) predicts that events (or subevents) with few alignments will be discarded.

Study 1. Study 1 asks how children compare events during verb learning when more naturalistic backgrounds are shown and some events are irrelevant. This is an important question because events that could be compared to each other usually are interleaved with other events. English-speaking 2 ½ year old children (range: 2;4-2;10), and 3 ½ year old children (range: 3;4-3;10), n=32 to date, participated in the study. Children were randomly assigned to one of 3 conditions. In the Relevant First condition (RRDDR), for a single verb, each child saw 2 events that show the same action, then two distracting events, then a relevant event before test (see Fig 1). In the Distractor First condition (DRDRR), each child first saw a distracting event, a relevant event, a distractor event and two relevant events before test. In an Alternating condition (RDRDR) no relevant events follow each other. This condition allows for the event immediately before test to be the same across conditions.

The event stimuli included a more complex background to better simulate the kinds of scenes children are viewing when they see events. Two sets of events were created in a kitchen and two took place in a park. In each, an animate agent causes a salient change in an inanimate patient, fitting prototypical causal event structure (Slobin, 1985).

The study began with 2 warm up pairs so children could practice pointing to the screen. Then the experimental trials started. When seeing a relevant event, children heard "Look! She's gonna <verb> it! She's <verb>ing it. She <verb>ed it." When seeing a distractor event, children heard "Now look at what she can do. Wow! Look! " during the event. These phrases were repeated until the child had seen 5 events. In the test phase, immediately before the test trial, children heard "Now it's your turn to find <novel verb>ing." and then saw two scenes in a split screen while hearing "Which one is <verb>ing it? Can you show me? Can you point to it? Where's she < verb>ing it?". In a second test trial, the same events were shown but on the opposite side of the screen as in Test Trial 1. The learning and test phase formed a single block of trials. The entire process was repeated until children had completed a block of trials for two novel verbs. Children's pointing behavior at test was coded as Correct or Incorrect or No Response.

A 2 (Age group: 2 ½, 3 ½) by 3 (Condition: Relevant First, Distractor First, Alternating) univariate ANOVA was computed; dv= correct extensions at test. A preliminary analysis reveals an effect of age, $F(1, 45) = 22.10$, $p < .001$, np2= .36. A one-sample t-test shows that 3-year-olds succeed across conditions, $t(22) = 12.81$, $p < .001$ (see Fig. 1). With additional participants, additional results may emerge.

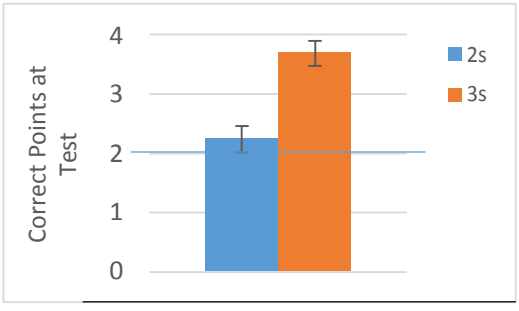


Fig. 1

Study 2.

Irrelevant events are oftentimes directly connected to relevant ones in everyday life. For example, a child may see her dad get a spoon from a drawer in the kitchen and

then turn and stir spaghetti. Previous work by Baldwin et al. (2001) and Hespos (2009) show infants can parse events embedded within a stream of action. If children are able to structurally align (or compare) similar segments of events, they should still be able to learn verbs even when the event referred to by the new verb is adjoined to a distracting event.

In Study 2, 2 ½- and 3½-year-old children (n= 30 to date) participated in one of two conditions. In the Relevant First condition (RD,RD,RD), for a single verb, each child saw a relevant event immediately adjoined to an irrelevant event, and then saw two more examples with that sequence. A similar logic was used for the Distractor First condition (DR,DR,DR).

The stimulus events were similar to those used in Study 1, except that the distractor event was filmed to appear immediately before the relevant event in one condition and immediately after the relevant event in the other condition. We also added difficulty to the test trials by creating a first pair of events using a new relevant event and a distractor event we had shown previously in the learning trials. In the second pair of events, children saw a new relevant and new distractor event (see Fig. 2).

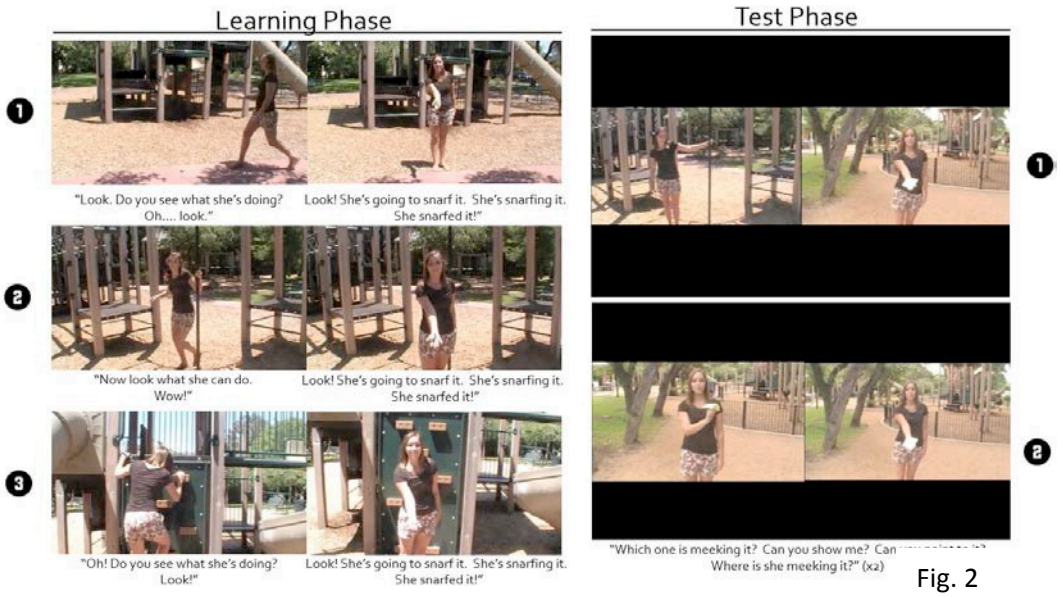


Fig. 2

The procedure was the same as that used in Study 1, except that distractor actions are attached to relevant actions (see Fig. 3).

The same analyses used in Study 2 as was used in Study 1. Preliminary results show a main effect of Age, $F(1, 39) = 21.47$, $p < .001$, with 3-year-old children's succeeding across conditions, $t(25) = 6.08$, $p < .001$ (see Fig. 4); additional participants are needed. Results from Loucks and Meltzoff (2013) suggest that preschoolers can reorder events based on the overarching goals of each set of across relevant examples, and thus, could be successful in all conditions, which would also be a useful result as no other verb study has (tested or) shown that children can disregard intervening events. Success in both studies despite intervening irrelevant events is consistent with predictions from structural alignment theory as irrelevant events or parts of events will not align with the target action.

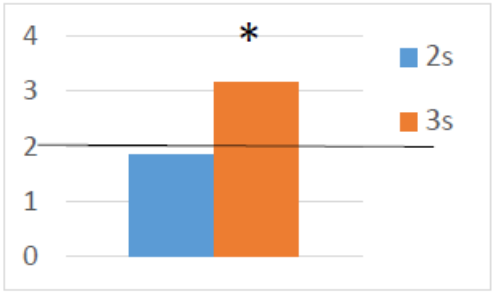


Fig. 3

EPISODIC AND SEMANTIC INFLUENCE IN OBJECT INSTANTIATION

Zachary Ekves (University of Connecticut), Gitte H. Joergensen (University of Connecticut), Xin Kang (The Chinese University of Hong Kong), Gerry T.M. Altmann (University of Connecticut)
zachary.ekves@uconn.edu

Throughout language processing one has to track object-state changes that occur during sentences and transform objects from their typical state to a changed state ("The chef will chop the onion…"). Subsequent to such changes, language can either refer back to that changed object where one needs to retrieve this changed object from episodic memory ("and then he will weigh the onion") or refer to a new object of that same type ("and then he will weigh another onion"). Previous research by Hindy et al. (2012) has shown that in cases where we refer back to these changed objects, there is increased activation in brain regions implicated in conflict and ambiguity resolution, which the authors argue arises from a need to select between the pre- and post-changed state of the object (a chopped or unchopped onion). However, this increased activation does not appear when one refers to "another" of that object type, suggesting that there is instantiation of a new object that is being selected as the referent (Solomon et al., 2015). Despite these results showing processing differences resulting from referring to old or new objects, little work has been done examining the effects of previous information on subsequent object representation. We conducted two eye-tracking experiments to investigate the representational content of these new object instantiations, specifically examining the role of object typicality and previous episodic information in this instantiation.

In the first experiment, we used a visual world paradigm to present participants with a series of potential object referents while they heard sentences that described situations in which an object underwent state change ("The woman will chop the onion"), and then that same object or another object of the same type was referred to ("Then, she will weigh the onion" or "Then, she will weigh another onion"). We examined looks to three critical object types in the scene (Figure 1): changed objects (e.g. a chopped white onion), unchanged objects (an intact white onion), and objects of the same type that were different kinds (an intact red onion). During the sentence final noun when we referred to "another" instantiation of the object ("…another onion") we observed significantly more looks to both the white and red intact onions than to the chopped white onion, with no difference between the two intact onions. We suggest that this indicates that "another onion" is interpreted as potentially referring to any onion (whether the white one or the red one) that is more typical than the atypical chopped onion. That is, "another onion" triggers the creation of a new discourse entity whose features are not inherited from the episodic memory of the original onion, but are instead inherited from semantic memory for more generic onions. The second experiment examined the extent to which previous episodic information influences subsequent representation of new object tokens (e.g., whether the "chopped-ness" of the first onion influences the featural instantiation of the new onion).

The second experiment used sentences of the same structure as experiment 1, but the items in the visual display differed (Figure 2) in one crucial way: The red onion – a relatively typical instantiation of an onion – was replaced by another chopped onion – a relatively less typical instantiation of an onion. Examining the influence of episodic feature information in this way requires there to be a specific chopped onion that is the referent in the first sentence. In order to effect this, the first sentence ("The chef will chop the onion") was played while only one object was displayed in the screen (one of the chopped white onions). During the second sentence ("and then he will weigh the/another onion") the additional critical objects, and distractors, were added to the scene. Critically, the intact onion in the display is the only (potential target) item that has the typical/generic features associated with onions (e.g., all of the chopped onions are relatively less typical exemplars of the onion category – essentially, they reflect a non-canonical form that onions can take). The addition of the second chopped onion allows one to instantiate a new object token that can share the same episodic characteristics as the initial object referred to in the first

sentence. If it's the case that the featural instantiation of new objects is completely isolated from the featural properties of the "old" object, we would expect to see more looks to the intact white onion than the second chopped onion. If it's the case that the episodic features of the original onion can leak into the featural instantiation of the new onion, we should see looks (albeit to a lesser extent perhaps) to the second chopped onion. In fact, during the final word "onion", there were no more looks towards the intact onion than towards the second chopped onion. However, by noun offset, there was a significant preference ($p = .002$) for the intact onion. The equivalent looks towards the two onions during "onion" suggest that any preference to instantiate "another onion" from semantic memory (i.e. as a typical, generic onion) was matched by an equivalent preference to instantiate "another onion" from episodic memory (i.e. with the same atypical feature set as the first, chopped, onion). Further analyses revealed that during "onion", there were more looks away from the second chopped onion than from the intact onion ($p = .001$). We interpret these patterns as suggesting that, initially at least, episodic features from the first (chopped) onion "leak over" to the instantiation of "another onion", causing participants to entertain the second chopped onion, albeit briefly, as a possible referent for the new onion. But this particular instantiation is not sufficiently compelling, and the eyes therefore move away. Thus, despite an influence of episodic information early on, there is an eventual settling on the new, more typical, object.

More generally, these results highlight the importance of previous episodic information acquired through event comprehension and the role of typicality in the representational instantiation of new objects. It further suggests a dynamic interplay between longer-term semantic knowledge (typicality information) and shorter-term episodic information that occurs in the representation of event participants as language unfolds.

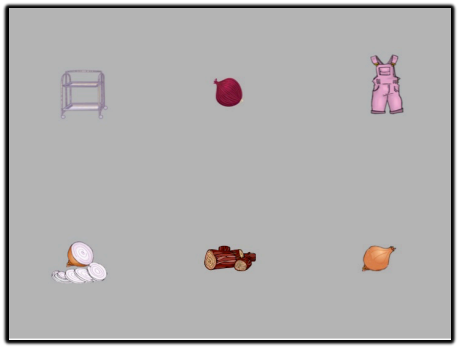


Figure 1: Visual scene for experiment 1.

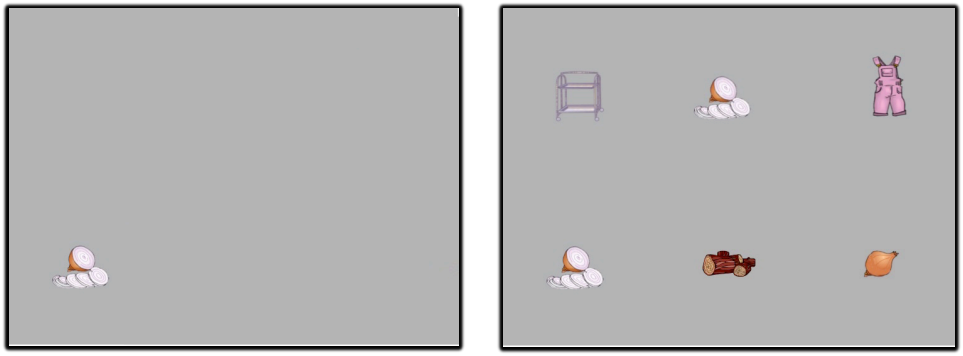


Figure 2: Initial and full display used in experiment 2.

Hindy, N. C., Altmann, G. T. M., Kalenik, E., & Thompson-Schill, S. L. (2012). The Effect of Object State-Changes on Event Processing: Do Objects Compete with Themselves? *The Journal of Neuroscience*, *32*(17), 5795–5803.

Solomon, S. H., Hindy, N. C., Altmann, G. T. M., & Thompson-Schill, S. L. (2015). Competition Between Mutually Exclusive Object States in Event Comprehension. *Journal of Cognitive Neuroscience, 27*(12), 2324-2338.

Can Infant Event Representations Give the Boot to Bootstrapping?

In this talk, I outline the basic assumptions and motivations behind bootstrapping proposals in language acquisition theory. I focus on semantic bootstrapping which was proposed by Pinker (1981) to solve the learnability problem of how the child identifies which words in the language are Nouns, which are Verbs, Adjectives, Prepositions, and so on. Pinker's answer to this problem was to propose that children initially assume a semantic correspondence between prototypical semantic types and syntactic categories: NOUN = person, place or thing; VERB = action; ADJECTIVE = description; PREPOSITION = location and so on. From these innate semantically based mappings, children then store distributional regularities associated with the category prototypes, and then generalize category membership based on distributional regularities associated with these semantic types.

I argue that such accounts treat the child language learner as a kind of code breaker. This is because it essentially treats language learning as trying to figure out how to map a set of preconceived categories on to form-meaning mappings arising from the language that children encounter in the linguistic input around them. This code-breaking approach assumes that children are not so much acquiring a grammar, but more or less mapping sounds onto some preconceived grammar specified within the genome. Unfortunately the genome does not look like something that could encode such specific knowledge, and strategies. On the other hand, theories of language acquisition do need to take seriously the problems of underdetermination and complexity that have been the hallmarks of arguments for nativist approaches to language acquisition.

My approach in this paper is to propose that the infant develops a rich set of representations that encode objects, events, spatial relations etc. in a highly structured manner. For example, their representation of events within the first year of life, prior to the onset of language production, includes the elements of argument structure that will serve as a basis for verb-argument structure when mapped onto language. The difference of this approach is that the infant is not trying to figure out the syntax of the linguistic input, but to map the syntax of the external language onto the syntax of the internal language, or what Fodor called the "language of thought." In this way, language learning is envisioned not as code breaking but, borrowing from Piaget, a process of assimilation (mapping the external language to the internal language), and accommodation (extending the internal representation to the wider conceptual scope of the external language.)

This proposal is accompanied by several kinds of evidence. The first is a summary of experiments I conducted with 6-12 month old infants in the prelinguistic stages of development. I briefly describe a series of experiments that provide evidence for argument structure in event representations of 6-12 month old pre-linguistic infants. Employing looking-

time habituation and eye tracking methods, these experiments show that, for 8-10 month olds, objects that are relevant to the meaning of an event (e.g., the transferred object in a GIVE event) lead to increased looking time when they are deleted from the event (GIVING with no object being transferred). However, when the same object does not play an argument role (HUGGING someone whilst carrying a toy), there is no increased looking time when the object is deleted. In numerous experiments using this paradigm, I demonstrate that infants only show dishabituation in situations where they have argument-structure representations, and that they have a conceptual understanding of the event itself. For example, at 10 months, while they appear to understand the act of GIVING, they do not understand the act of SHOWING, perhaps because the latter requires aspects of theory of mind that are not yet developed.

I argue that the initial mapping of internal to external language in raw form produces a core language that includes the basics of what is seen in home-sign systems and certain aspects of pidgin-like languages. Missing in this core language are the fragile properties of linguistic encoding that are language specific such as temporal and aspectual coding of events, incorporation of theory of mind elements in language (e.g., structures encoding mental states and propositional attitudes.) In addition, core language offers only primitive elements of domain knowledge like number (small number exact knowledge plus large number estimation), and color (discrimination but not conceptual representation.) In such cases, it is the external linguistic system that drives development of the internal language (i.e., accommodation).

Within this framework, the goal of language learning is not to figure out which words are nouns and verbs etc., since this would simply emerge from the process of mapping the internal language to the external language. Giving a label to a category is not necessary because the assimilation process is itself an embodiment of those categories within the basic event representations that the infant brings to language learning.

Preference for single events guides perception in Russian: A phoneme restoration study.

How do language users associate event representations with syntactic strings, when those strings often underdetermine event construal? One possibility is that the processor builds a single event unless the context or grammar suggests otherwise. While there are many possible origins for a single event preference (e.g., cognitive or representational simplicity) evidence for such a preference has been observed in areas as diverse as collective/distributive ambiguities (Clifton & Frazier, 2012), coordination (Frazier & Clifton, 2013, also Hoeks et al., 2002), quantifier domain ambiguities (Harris et al., 2013), and reciprocals (Fiengo & Lasnik, 1973; Majewski, 2014). In this study, we present evidence that the perception of complex predicates in Russian is similarly influenced by a general preference for a single event construal. In essence, perceivers are biased to resolve distorted speech towards a single event interpretation when grammatically licensed.

Serial coverb constructions (see also auxiliary verb constructions) are generally defined as a monoclausal verb complex that shares its arguments and inflectional features like tense and aspect, and are interpreted as single events (Anderson, 2006; Aikhenvald, 2011; Bisang, 2009). Serialization appears in a whole host of languages worldwide, and often alternates with a conjunction, e.g., English *go (and) listen to the radio*. Typically, a restricted set of motion verbs acquires a semantically bleached usage through grammaticalization, though they usually retain both this light and a fully lexicalized use (Traugott & Dasher, 2001; Anderson, 2006; Butt, 2010). Russian is no exception, as a bleached auxiliary use of light verbs of motion (e.g., *idu* 'I walk/go') may immediately precede another verb (*idu slushaju* 'I walk/go listen'), giving rise to unambiguous single event interpretations. Like English and other Germanic languages, an intervening conjunction (e.g., *idu i slushaju* 'I walk/go and listen') is ambiguous between a single and a multiple event interpretation. In the former, the 'going' event is a subpart of the listening event, in which one may initiate an action that results in listening. In the latter, going and listening describe distinct events.

We predicted that when listeners were tasked with resolving an ambiguous or uncertain acoustic signal, they would choose the option which allows a single event structure whenever permitted by the grammar. To test this prediction, we used the phonemic restoration method, in which an intruding cough, tone, or other noise masks an excised phoneme in the acoustic signal, allowing subjects to perceptually "restore" the missing phoneme (Warren, 1970), especially in highly biasing or constraining contexts (Warren & Sherman, 1974; Samuels, 1981). Recent research has shown that perceivers use high-level contextual information in restoration decisions. For example, Stoyneshka et al. (2010) showed that perceivers used prosody in determining how to resolve a masked segment in cases of temporary ambiguity (also Carbary et al., 2015). Another case is Mack et al. (2012), who find that subjects restored a zero subject, e.g., *It* in *(It) seems like it's going well*, when context supported the pragmatic usage preferences for doing so.

We measured the rate at which Russian speakers restore a conjunction (a single phoneme 'и' *i*) between two verbs, manipulating verb type (Light, Lexical) and Adverb status (Present, Absent) in a crossed 2x2 design (1). A single event preference predicts that subjects will restore a conjunction after a Light verb less often than after a Lexical verb, which necessitates two events, unless an Adverbial indicates a separate clause, precluding a serial verb construction.

Twenty-seven native speakers of Russian listened to Russian sentences that were obscured with pink noise in selected locations over headphones in a sound attenuated anechoic chamber. Subjects repeated into a microphone what they thought the message was before it had

been distorted, and their responses were recorded. The critical sentences masked * the area between verb phrases, as in (1). As an Internet pilot study (N = 16) revealed that phonetic cues for the conjunction were acoustically present as coarticulation on preceding vowels (also Choi & Keating, 1991), no vowels were included before the conjunction site. Further, materials were equally balanced between cases in which the conjunction was and was not produced in the original recording. All materials were created and produced by a native Russian speaker. In addition to filler sentences, twelve control sentences with narrowly constrained grammatical options for restoration (e.g., case marking) were included to assess speaker competence and attention to the task. Five subjects were removed for failing to accurately restore control items.

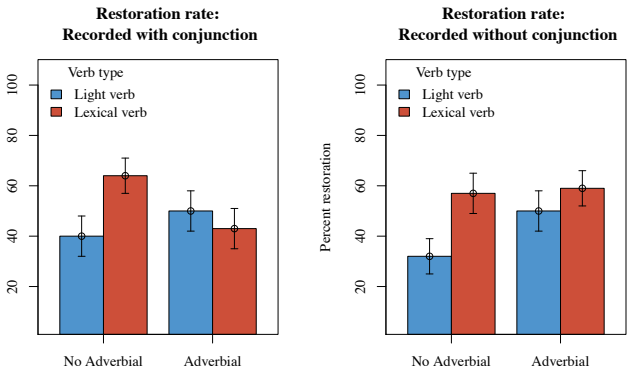
The rate of restoration of a conjunction between verbs was analyzed as a linear mixed effects logistic regression model with sum-coded predictors. Subjects were less likely to restore a conjunction after an initial Light verb (1a) than an initial Lexical verb (1b), t = -2.37. However, this difference only affected cases in which serial verbs, and by hypothesis single events, were permitted: subjects restored a conjunction for Lexical verbs more often than for single Light verbs (d = 24%), but not when an Adverbial intervened (d = 2%), t = -2.10. Models that included whether the item was originally produced with a conjunction did not change the results (Fig. 1). An additional post hoc analysis on non-restored items revealed that subjects produced non-significantly longer pause durations after Adverbials following a Light verb, $t = 1.6$, raising the possibility that subjects tended to add a prosodic break to signal a clause boundary in such cases.

In sum, the rate of conjunction restoration between verbs, and thus perception of an ambiguous or noisy signal, was guided by a general preference for single event interpretation when grammatical possible. We propose that the bias for single events indicates a preference for representational economy of complex situations in the discourse model, which interacts with grammatical and perceptual decisions. Our results are thus compatible with growing evidence that conceptual constraints restrict how the processor structures bottom-up information, especially for ambiguous input (e.g., Stoyneshka et al., 2010).

(1) **Sample materials.**
a. *Light verb (Adverbial).*
Идем (не торопясь) * слушаем радио.
*idjom (ne toropjasja) * slushaem radio*
*walk (neg haste) * listen radio*
'We walk/go (slowly)* listen to the radio.'

b. *Lexical verb (Adverbial).*
Едем (не торопясь) * слушаем радио.
*edem (ne toropjasja) * slushaem radio*
*drive (neg haste) * listen radio*
'We drive (slowly) * listen to the radio.'

Fig. 1. Conjunction restoration rate; item originally produced with and without conjunction.



PRAGMATIC INTERPRETATION OF INFORMATIONALLY REDUNDANT EVENT MENTIONS

Ekaterina Kravtchenko (Saarland University), Vera Demberg (Saarland University)
eskrav@coli.uni-saarland.de

Work in pragmatics shows that speakers typically avoid stating information already given in the discourse [1]. However, there has been little investigation of how comprehenders interpret utterances which assert what can easily be inferred using prior knowledge. Previous work [2] has shown that informationally redundant event descriptions trigger context-dependent implicatures, which increase utterance utility in line with listener expectations [1,3]. In a replication and 3 follow-up experiments, we look at utterances which refer to event sequences (*scripts*, such as *going to a grocery store*), and find that modifications to the informational status of the redundant event utterance significantly change the likelihood with which a 'compensatory' inference is drawn.

Design: 24 items; 2 (typical vs. atypical context) x 2 ('predictable' vs. optional event). 'Predictable' events (4a) were directly implied by the (typical) script, while optional events (4b) were not. Initial context was typical (1a), or implied the 'predictable' event was optional or atypical (1b).

Typical context

[1a] John often *goes to his local supermarket, as it's close by*$_{\text{typical}}$.

Atypical context

[1b] John often *doesn't pay at the local supermarket, as he's usually broke*$_{\text{atypical}}$.

[2] Today he entered the apartment with his shopping bags flowing over. He ran into Susan, his best friend, and talked to her about his trip. Susan then wandered over to Peter, their roommate, who was in a different room.

[3] She commented: "John went shopping.

EXP 1-2 [4a/b] He {***paid the cashier***$_{\text{a-predictable}}$ | *got some apples*$_{\text{b-optional}}$} {!$_{\text{exp1}}$ | .$_{\text{exp2}}$}

EXP 3 [4a/b] Oh yeah, and he {***paid the cashier***$_{\text{a-predictable}}$ | *got some apples*$_{\text{b-optional}}$}.

EXP 4 [4a/b] You'd be surprised – he {***paid the cashier***$_{\text{a-predictable}}$ | *got some apples*$_{\text{b-optional}}$}.

[5] I just saw him in the living room."

Q: How often do you think John usually {*pays the cashier* | *gets apples*}, at the store?

Procedure: Participants (all N=400, MTurk) see 6 of 24 stories; 4 show both context and the critical utterance (*posterior* measure; [1-5] in the example item on p. 2), and 2 only context (*prior* measure; [1-2]). After reading, participants rate the perceived 'typicality' of events on a continuous scale. We used a between-subjects design, where differences between *prior* and *posterior* ratings were examined across, rather than within, participants.

Experiment 1: This experiment replicated the results in [2], which used a within-subjects design. An LME model showed a *context* by *event predictability* interaction (β=-14.61, p$<$.001) for predictable events, indicating that comprehenders assign informationally redundant event mentions (4a: 'John went to the store. He paid the cashier!') an 'informative' interpretation, by reinterpreting the 'predictable' activity as atypical in context (i.e. 'John does not typically pay the cashier'). No (re-)interpretation occurs for event mentions that are informative a priori (where the script does not automatically imply the event occurred) – or where the scenario is changed to make the typically inferable event uninferable.

Experiment 2: In Exp. 1, the event utterance was followed by an exclamation point, which signals importance of the information, as well as speaker intentionality in conveying it. In Exp. 2, we replaced the exclamation point with a period, to determine whether the inference is dependent upon (implicit) intonational support. Given the general abnormality of informational redundancy [4], we expected an effect similar to Exp. 1. Design and procedure was otherwise identical. The results show, contrary to expectations, that while a significant interaction remains, redundant event mentions no longer give rise to any 'atypicality' inferences (β=-0.13, n.s.). A subgroup analysis of 'predictable' events suggests that very highly inferable events (those rated the most highly 'typical,' given just context) are still interpreted as atypical (β=-3.01, p$<$.05) when mentioned, while less highly inferable events are not (β=17.03, p$<$.01). This shows that informational status, independent of the literal content or informativity of an utterance, significantly influences the likelihood of it giving rise to an inference *based on its informativity* (specifically, event typicality/inferability).

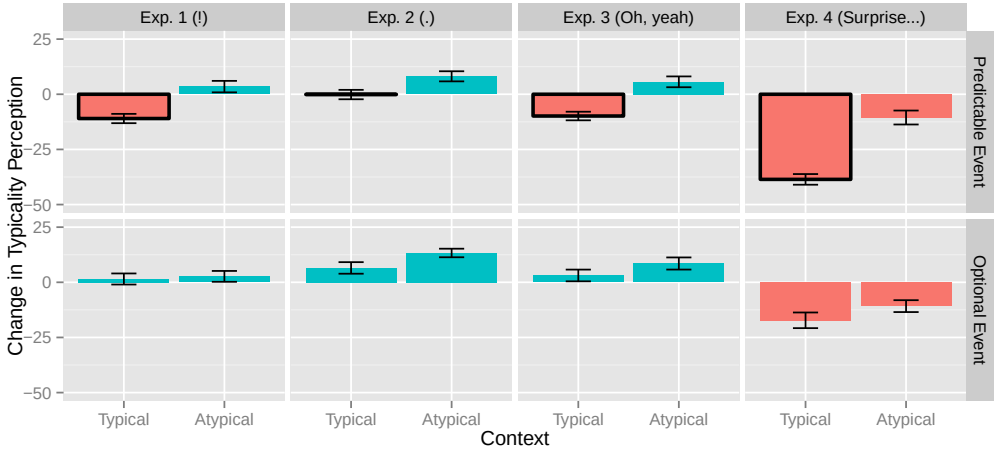
Experiment 3: Based on the results of Exp. 2, we hypothesized that linguistic markers not explicitly associated with surprisal, but indicating intentionality and relevance of the information conveyed, would produce effects of accommodation to informationally redundant utterances similar to Exp. 1. Information structure was manipulated by introducing the event utterances with *oh yeah, and....* This implicitly indicates speaker intentionality in mentioning the event, as well as the speaker's belief that it may describe something of interest/relevance to the listener. It was predicted that, overall, the results should mirror that of Exp. 1, with comprehenders assigning redundant event mentions 'informative' interpretations. A quantitatively and qualitatively similar interaction between *context* and *event predictability* was in fact found (β=-15.44, p$<$.001), with a very similar pattern of results.

Experiment 4: Based on Exp. 1, and to provide a control measure, we also hypothesized that explicit markers of surprisal, in contrast to an exclamation point (which does not directly convey surprisal), would produce more dramatic effects – inducing strong 'atypicality' inferences in all conditions, most significantly in the informationally redundant condition. Information structure was manipulated by inserting phrases that explicitly signaled surprise (e.g. *you'd be surprised, but...*). As predicted, significant effects were found in all conditions, with a significant *context* by *event predictability* interaction for predictable events (β=-28.04, p$<$.001).

Overall, these findings are consistent with the claim that mention of easily inferable material is systematically reconciled with the assumption of an informative speaker [4], giving rise to context-dependent implicatures regarding event 'typicality.' The results from Exp. 2 show, however, that this effect is highly modulated by informational status, or at least dependent on implicit contextual support for the specific inference. Additionally, it shows effects of implicit prosody on implicature generation that are different from better-known effects of prosodic focus. Results from Exp. 3-4 show that other changes in informational status, which explicitly signal surprise or utterance relevance/intentionality, replicate the original effect.

All experiments provide robust evidence that excessive informational redundancy is perceived as anomalous, and that comprehenders alter their situation models in order to accommodate it. However, there is also a substantial effect of informational status on implicature generation, despite the implicature being based on the prior inferability of the event in question - which does not change across experiments. This supports, in part, Grice's observation [4] that overinformativity, unlike underinformativity, despite often being treated similarly in later literature, is not necessarily problematic for listeners; while also complementing work in the dialogue literature [5], which shows that informationally redundant utterances are both common, and frequently used to convey 'informative' non-literal content.

Figure 1: Bars show relative changes in perception of activity typicality. Positive values indicate increases in perception of activity typicality; negative values indicate decreases. Informationally redundant conditions are highlighted.



[1] Horn. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. **[2]** Kravtchenko & Demberg. (2015). Semantically underinformative utterances trigger pragmatic inferences. **[3]** Atlas & Levinson. (1981). It-clefts, informativeness and logical form: Radical pragmatics. **[4]** Grice. (1975). Logic and conversation. **[5]** Walker. (1993). Informational Redundancy and Resource Bounds in Dialogue.

MICROSENSES: EXPLORING MULTI-WORD EVENT CATEGORIZATION

Aron Marvel (University at Buffalo), Jean-Pierre Koenig (University at Buffalo)
aronmarv@buffalo.edu

One function of verbs is to name event categories. But event categories can receive their "names" from the combination of several words. An event category named by a single verb in one language can be conventionally coded via multi-word expressions in others (compare Spanish *entrar* and English *go in*). Furthermore, sentences with the same verb used in the same sense, according to best lexicographic practice, often belong to different event categories. Compare a) *The senator raised a glass in celebration* and b) *The crane raised the car out of the water*. While (a) describes a toast, (b) describes the extraction of a large object. Because the verb sense is unchanged, the information relevant for event categorization must come from sources outside of the verb and these sources combine with the verb to form complex event names.

Two questions motivate the research presented here. First, to what degree can native speakers recognize event categories that go beyond verb senses? Second, how can linguistic resources outside the verb serve as cues in making such distinctions? We explore answers to these questions in three ways. A rater study contrasts event categories with verb senses and gauges the relative contribution of subjects and objects as cues for categorization. A native speaker judgment task is then used to validate the results of the rater study and offers a more fine-grained representation of event similarity on a Likert scale. And lastly, subject and object meanings are tested as cues for event categorization in a computational model using Latent Semantic Indexing (LSI) and hierarchical clustering.

Based on previous research, we propose six parameters for distinguishing event categories beyond verb senses. They are listed in Table 1. These parameters were used as guidelines to sort into event categoriesa pseudo-random sample of 200 sentences for each of ten verbs *bake, borrow, buy, cover, deliver, frighten, immerse, pull, rescue, sell* from the British National Corpus (BNC). For simplicity, only verb plus subject and verb plus direct object combinations were analyzed. Results revealed an average 16.5 event categories per verb, in comparison to 3.8 senses according to the American Heritage Dictionary. 62% of event categories were distinguished by verb plus direct object combinations. This asymmetry was reversed for the object-experiencer verb *frighten*, suggesting the differential contribution comes from differences in semantic properties of the event participant types that typically occur in subject and direct object positions.

Results from the rater study were validated in an experiment where 30 participants were presented pairs of sentences from our corpus and asked to judge the similarity of the described events. Two of our ten verbs (*bake, immerse*) were excluded due to relatively sparse corpus data. The remaining eight verbs were used to construct 96 sentence pairs, balanced across three groups: in the first group each sentence pair shared both verb sense and event category according to our parameters; in the second group each pair differed in both sense and category but shared the same verb; in the third group each pair shared a verb sense but differed in category. The median similarity rating for each participant was used as a breakpoint and responses compared against same/different event category judgments from the rater study in a Chi-square test of independence. The results strongly suggest a relationship between similarity and inclusion of events within the same category, with 72% of similarity ratings in accordance with category judgments from the rater study and a medium to large effect size as measured by Cramer's V (X^2=218.64, N=1129, p<.001, V=.44). The results are shown in Figure 1.

The categorization and similarity judgment results suggest that event categories are distinguished at a more fine-grained level than that of the verb sense (at a ratio of at least 4:1), and that the contribution of participants named by subjects and objects in making that distinction is not symmetrical. But the results do not give us a way of quantifying the knowledge that *raising a car* is much more like *raising a truck* than either is like *raising a glass*. To quantify this knowledge, Latent Semantic Indexing (LSI) and a semantic space constructed from the BNC

were used to obtain relatedness values for subjects and direct objects. A mixed effects logit model suggested that LSI event participant relatedness values are reliable predictors of category inclusion, with higher relatedness values indicating greater odds of sharing an event category (S β=179.14, z=19.87, p<.001, DO β=69.87, z=12.03, p<.001). LSI relatedness values were also a marginally significant predictor of the odds of an item's being above a participant's median similarity rating (subject β=2.31, z=2.07, p=.04; direct object β=-1.70, z=1.58, p=.12).

The relationship between LSI relatedness and category inclusion allowed us to develop and test a computational model of event categorization. Pair-wise LSI relatedness values were obtained for each verb's subjects and separately its direct objects. Values were clustered using average linkage, and a cross-section of the clusters was taken when the number of clusters was equivalent to the number of the rater study event categories for that list. The resulting cluster categories were compared against the rater study categories, using the latter as a Gold Standard. An F value was calculated from the harmonic mean of the precision and recall measurements for the cluster categories. As a measure of comparison, each list was assigned 100 iterations of random categories and the resulting F value compared to the F obtained from LSI relatedness. On average, using LSI relatedness values gave a 42% improvement over random categories.

Research in semantics has shown event categories are not always named by single verbs. In this paper, we explore parameters for distinguishing event categories beyond single verb senses. Our results suggest (1) that raters can distinguish four times more event categories than dictionaries recognize verb senses, (2) some types of participants contribute more than others to distinctions among event categories and (3) similarities between event participants are good cues for distinguishing among event categories that go beyond single verb senses.

Agent type: Do the events involve agents of a distinct number, animacy, abstractness? E.g. *the fleet rescued* vs. *the medic rescued*
Time frame: Do the events last longer/shorter? E.g. *built a sand castle* vs. *built a skyscraper*
Sociocultural salience: Do the events play a special role in the practices of a community? E.g. *borrowed a book* vs. *borrowed an iron*
Available inferences: Is there additional pragmatic/semantic information obtained from world knowledge about typical event scenarios that the events take part in? E.g. *covered her bruise* vs. *covered her eyes*
Complexity: Do the events include particular subevents as well as relations among sub-events and their participants? E.g. *bought a house* vs. *bought sunscreen*
Specific motion sequence: Are actions performed according to a recognizable motor program? E.g. *pulled the ball* (soccer) vs. *pulled the cart*

Table 1*. Parameters relevant for event categorization*

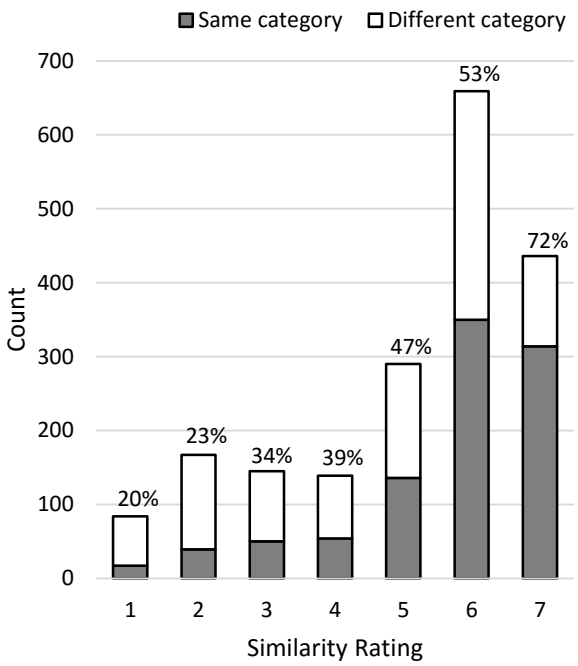


Figure 1. *Count and percentage in same category for each similarity rating (includes only pairs which share a verb sense)*

WHAT YOU('D) SAY IS WHAT YOU FOCUS ON:
LANGUAGE EFFECTS ON MOTION EVENT CONCEPTUALIZATION

Guillermo Montero-Melis (Stockholm University)
montero-melis@biling.su.se

How much is event cognition guided by language? It is well established that speakers of different languages describe motion events in different ways, especially with respect to how path and manner information is syntactically realized (Talmy, 2000). However, the issue of whether these linguistic differences affect nonverbal event conceptualization remains unsettled (e.g., Papafragou, Massey, & Gleitman, 2002 for no language effects; Gennari, Sloman, Malt, & Fitch, 2002; Papafragou, Hulbert, & Trueswell, 2008; Papafragou & Selimis, 2010 for effects only under certain conditions; and Kersten et al., 2010; Lai, Rodriguez, & Narasimhan, 2014 for clear language effects). Moreover, motion events can be described in different ways within the same language. Will short-term exposure to contrasting ways of describing events result in similar contrasts in how they are conceptualized nonverbally?

A series of studies explored the connection between how events of caused motion are described and how similar these events are perceived to be, focusing on Spanish and Swedish, two languages that lexicalize motion in contrasting ways. Caused motion (e.g., 'the boy rolled the tyre into the barn') adds a layer of complexity in event structure compared to spontaneous motion (e.g., 'the boy ran into the barn') and has received less attention in the Talmy-inspired literature. The stimulus set consisted of 32 video animations that systematically cross different paths (e.g. into, up) with different manners of manipulating objects (e.g. rolling, dragging). To test how speakers judge event similarity we implemented a novel similarity arrangement paradigm (based on Goldstone, 1994), see Figure 1.

Figure 1. Example of final arrangement in nonverbal task. Stills represent events. Similarity is based on pixel distance.

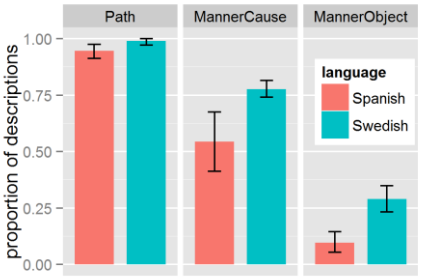


Figure 2. Study1: Proportion of descriptions mentioning path (e.g. into), manner of cause (e.g. push) and manner of object (e.g. roll). Error bars = by-subject 95% C.I..

Study 1 first sought to establish whether Spanish (N=18) and Swedish (N=19) speakers differ in how they *describe* caused motion events. In line with previous findings on spontaneous motion, Swedish speakers were more likely to include information about manner, such as whether the agent pushed the object (manner of cause) or whether the object rolled (manner of object); speakers of both languages were equally likely to include path information (Figure 2). Do these linguistic differences correlate with nonverbal event conceptualization?

Study 2 probed whether Spanish (N=47) and Swedish (N=47) speakers perceive event similarity differently. Participants were randomly assigned to one of two conditions. In the *linguistic* encoding condition participants described the events prior to carrying out the nonverbal task, while in the *free* encoding condition they were simply familiarized with the events before the nonverbal task. In both conditions Swedish speakers were more likely than Spanish speakers to base their similarity arrangements on the manner of object motion (rolling/sliding). There was no significant difference in how much speakers of either language relied on the manner of cause (push/pull) or on the path of motion (e.g., into or up), see Figure 3. Interestingly, in the linguistic encoding condition participants of both languages were more likely to rely on the path of motion (left panels Figure 3). That is, describing the events

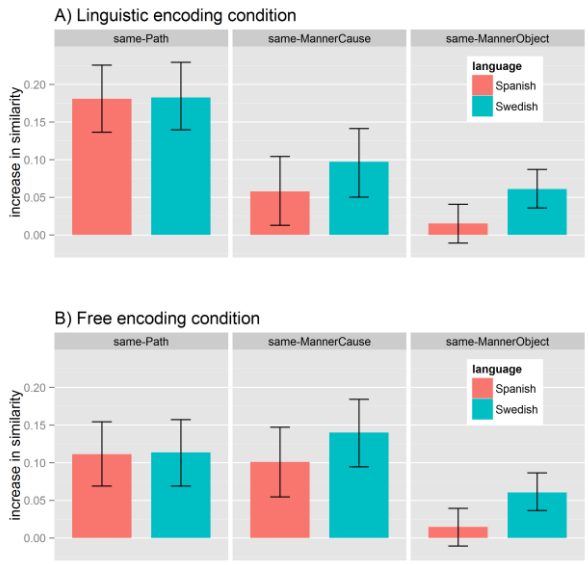


Figure 3 . Study2: Increase in similarity due to shared path, manner of cause and manner of object in each language, by language (colour) and encoding condition (upper/lower figure). Error bars = 95% C.I.

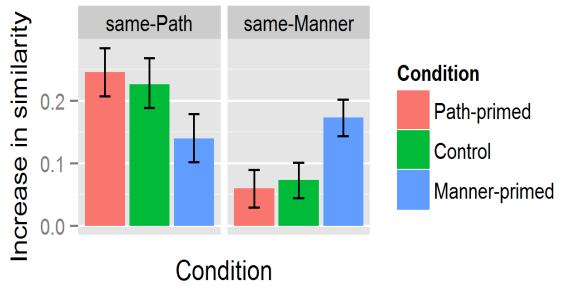


Figure 4 . Study3: Increase in similarity due to shared path and manner by priming cond. Error bars = 1 SE.

prior to judging their similarity had the same effect on Spanish and Swedish speakers. Is this correlation between native language and conceptualization resistant to linguistic manipulation in a second language (L2)?

Study 3 tested whether native Swedish learners of L2 Spanish (N=60) could be led to rely on different event components in event similarity judgments, by manipulating their recent L2 experience. Three between-subject conditions manipulated whether participants were path-primed in L2, manner-primed in L2, or did not receive any priming, during an encoding phase in which they described animations of caused motion events in their L2. In the subsequent test phase, participants engaged in the same nonverbal task as in Study 2. Manner-primed participants relied significantly more on the manner dimension than path-primed participants (Figure 4 right panel), but they only showed a trend towards relying less on the path dimension (Figure 4 left panel). Participants that had not been primed patterned overall like path-primed participants (Figure 4 middle bars).

Together these results contribute with several insights to the relation between linguistic expression and nonverbal event conceptualization:

1. Speakers of different languages may vary in how they linguistically encode events of caused motion (Study 1).
2. Cross-linguistic differences partly correlate with how event similarity is perceived: Swedish speakers pay more attention than Spanish speakers to the object manner of motion, but both rely equally on agent manner and path (Study 2).
3. Describing the events before judging their similarity (as opposed to just watching them) makes the path dimension more salient in both languages (Study 2).
4. Effects of language on event similarity perception can be modulated by linguistic priming, suggesting that event conceptualization relies on *ad hoc* categories (Study 3).
5. Path seems to be a more central feature than manner, both in descriptions and in similarity judgments (Studies 1–3).

This suggests that the way events are usually described in one's language partly affects how one perceives them – what you('d) say is what you focus on. This is because linguistic categories map onto conceptual categories that can be recruited during tasks that do not overtly involve language. While this mapping is influenced by our native language, it is also flexible and can be modulated by recent linguistic experience that clashes with habitual lexicalization patterns. Conversely, preliminary analyses also suggest an effect in the opposite direction: carrying out a nonverbal similarity task seems to affect subsequent event descriptions (what you focus on is what you say).

Seeing events in a second a language: A cognitive "accent" in event perception

Irina Potapova (San Diego State University, University of California, San Diego),
Lera Boroditsky (University of California, San Diego)
ipotapova@ucsd.edu

How do the languages we speak shape the way we construe and perceive events? We consider event perception in bilinguals and ask what role linguistic habits in your first language play in how you construe events in a second language. We focus on grammatical aspect, and compare Russian-English bilinguals (tested in English) with English monolinguals.

Aspect is obligatorily marked in Russian. To say "he ate cereal," Russian speakers must choose either a perfective form of "ate" (*s'yel*, *doyel*) indicating a completed action, or an imperfective form (*yel*, *poyel*, *nedoyel*) to indicate an event in progress or an incomplete event (1). Unlike the English "ate," there is no grammatical form available in Russian that could apply both for a complete and an incomplete event. While completion information can be supplied in English in a variety of other ways, this is not obligatory. Does this difference between the languages lead to differences in the cognitive habits of Russian and English speakers?

Prior work has shown that differences in aspect marking between languages indeed predict differences in how people construe events (2). Further, bilinguals' event similarity judgments reveal influences from their second language when tested in their native language (3, 4). In this paper we ask about attentional patterns that might persevere as bilinguals shift from a language that makes more obligatory distinctions to one that has fewer. That is, do bilinguals construe events with a cognitive accent imparted by the requirements of their first language?

In Study 1 we ask: when Russian-English bilinguals speak English, do they include additional information about completion (over and above what is included by monolingual English speakers)? In Study 2 we ask: when making non-linguistic judgments about events, do Russian-English bilinguals attend more to completion than do monolingual English speakers?

Study 1: 43 English monolinguals (18 males; Mage = 32.56 yrs; SD = 8.55) and 38 Russian-English bilinguals (11 males; Mage = 32.97 yrs, SD = 8.41) participated. Each participant viewed 12 events (six complete, six incomplete), with each event depicted by two photos shown sequentially. An example event showed a man eating cereal in image 1, then the same man talking on the phone with the cereal bowl sitting on the table in image 2. The complete version of this event showed an empty cereal bowl in image 2. In the incomplete version, there was some cereal still left in the bowl, as if the man had been interrupted by the phone call. Each participant saw either the complete or the incomplete version of each event.

After seeing the two images, participants were given a sentence frame (e.g., Yesterday, he _____ cereal) and a verb (e.g., "TO EAT") and were asked to fill in the blank using the verb to provide a good description of the event. The blank was left long to allow for (and invite) elaborated descriptions, and participants were explicitly instructed that they were welcome to include additional words. Each response was coded as either providing explicit (in)completion information (e.g., *he ate all of the cereal; he ate some cereal*) or not (e.g., *he ate cereal*).

English monolinguals used simple past tense to describe both complete and incomplete events on more than 75% of the trials. Russian-English bilinguals provided more (in)completion information than English monolinguals for incomplete events, as revealed by a significant interaction in a 2x2 (group x event type) ANOVA, $F(1, 79) = 4.406$, $p = .039$. Of the bilinguals, 23 were sequential bilinguals with Russian as their first language. This set provided more (in)completion information than English monolinguals for both complete and incomplete events, $F(1, 64) = 12.239$, $p = .001$ (see Fig. 1). For incomplete events, these bilinguals included (in)completion information (48.7%) nearly twice as often as English monolinguals (24.57%).

Study 2: A new set of 33 English monolinguals (23 males; Mage = 32.27 yrs, SD = 8.53) and 34 Russian-English bilinguals (8 males; Mage = 35.03 yrs, SD = 11.80) participated. On each trial, participants saw a reference event and two comparison events. Their job was to decide which

of the two comparison events was most similar to the reference event. In an example trial, the reference event might show the complete version of a man eating cereal. One comparison would show the incomplete version of the same man drawing a tree (an agent match). The other comparison event would show the complete version of a different man cleaning the table (a completion match). The events were shown entirely in pictures, and participants clicked on their choice, making non-verbal responses. Participants were never asked to describe the events.

As predicted, Russian-English bilinguals selected the completion match much more frequently (65.15%) than monolingual English speakers (25%), $F(1, 65) = 24.978$, $p < .001$ (see Figure 2). Performance was consistent for both sequential and simultaneous bilinguals.

Discussion: These findings suggest that Russian-English bilinguals differ from English monolinguals in both how they describe and perceive past events. Despite completing the event description task in English, individuals with experience speaking Russian voluntarily provided additional information about completion (particularly when viewing incomplete actions) over and above what was provided by English monolinguals. The bilinguals were also more sensitive to completion information in a non-linguistic task. While English monolinguals preferred to match events based on who was completing the action, Russian-English bilinguals matched events based on degree of completion. To them, completion information was more salient, even in a non-linguistic task (with instructions provided in English). It appears that experience with obligatory grammatical aspect in the bilinguals' native Russian manifests in both a semantic accent (5) in a productive language task and a cognitive accent (6) in a non-linguistic task.

Figure 1

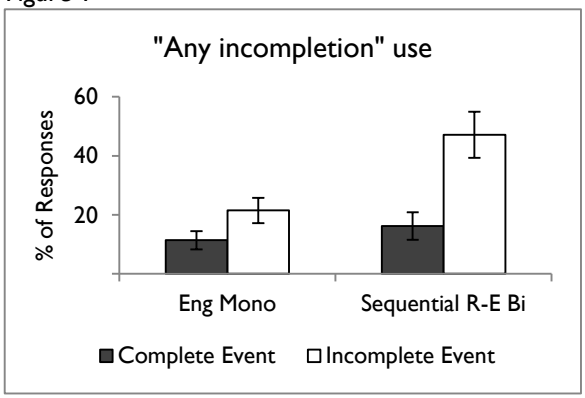
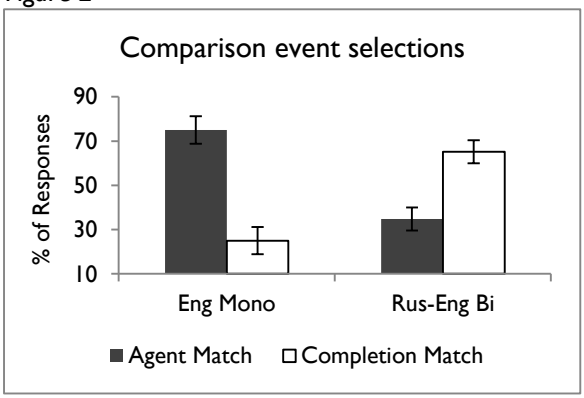


Figure 2



References

1. Ramchand, G. C. (2008). Perfectivity as aspectual definiteness: Time and the event in Russian. *Lingua, 118(11)*, 1690–1715.

2. Athanasopoulos, P., & Bylund, E. (2013). Does grammatical aspect affect motion event cognition? A Cross-Linguistic comparison of English and Swedish speakers. *Cognitive Science*, 37(2), 286–309.

3. Athanasopoulos, P., Damjanovic, L., Burnand, J., & Bylund, E. (2015). Learning to think in a second language: Effects of proficiency and length of exposure in English learners of German. *The Modern Language Journal*, *99*(S1), 138-153.

4. Bylund, E., & Jarvis, S. (2010). L2 effects on L1 event conceptualization. *Bilingualism: Language and Cognition, 14(01)*, 47–59.

5. Lucy, J. (2004). Language, culture, and mind in comparative perspective. *Language, Culture, and Mind*, 1–22.

6. Bylund, E. & Athanasopoulos, P. (2014). Linguistic relativity in SLA: Towards a new research programme. *Language Learning*, 64(4), 952-985.

READING BETWEEN THE LINES: THE INFLUENCE OF SCRIPT KNOWLEDGE ON ON-LINE COMPREHENSION

Elisabeth Rabs, Heiner Drenhaus, Francesca Delogu, Matthew Crocker (Saarland University)
erabs@coli.uni-saarland.de

While the influence of linguistic context on language processing has been extensively studied [1,2], less is known about the mental representation, structure and use of so-called *script* knowledge. Scripts [3,4] are defined as a person's knowledge about temporally and causally ordered sequences of events. They are often activated by linguistic context, but otherwise left implicit. In two ERP studies we examine how such non-linguistic event knowledge influences predictive language processing (as addressed by e. g. [5,6]) beyond what linguistic prediction or lexical priming alone can explain. Specifically, we find evidence for a decrease in N400 amplitude - known to reflect a word's unexpectedness [7] - for target nouns consistent with events that are expected according to script knowledge.

Experiment 1 focuses on differentiating the relative contribution of lexical priming and script knowledge. Assuming the temporal structure of scripts is accessible and used for prediction, but does not alter any influence of priming, we inserted temporal shifts affecting the plausibility of the critical object.

Intro:
Am Abend ging Peter zum Fondue essen.
(In the evening Peter went to eat a fondue.)

Target sentence:

(1) Minor temporal shift / Script-fitting target	Einen Moment später schmolz er den *Käse*. (A moment later he melted the *cheese*.)
(2) Minor temporal shift / Neutral target	Einen Moment später schmolz er den *Schnee*. (A moment later he melted the *snow*.)
(3) Major temporal shift / Script-fitting target	Einen Tag später schmolz er den *Käse*. (One day later he melted the *cheese*.)
(4) Major temporal shift / Neutral target	Einen Tag später schmolz er den *Schnee*. (One day later he melted the *snow*.)

Participants were presented with a context sentence introducing a script. The context sentence was followed by a transitive target sentence, presented word-by-word, in which the direct object resulted in either a script-fitting event (1/3) or a neutral event (2/4). Additionally, target sentences started with a temporal adverbial indicating that either the script mentioned before was still active, expressed by a minor temporal shift (1/2), or likely inactive, expressed by a major temporal shift (3/4).

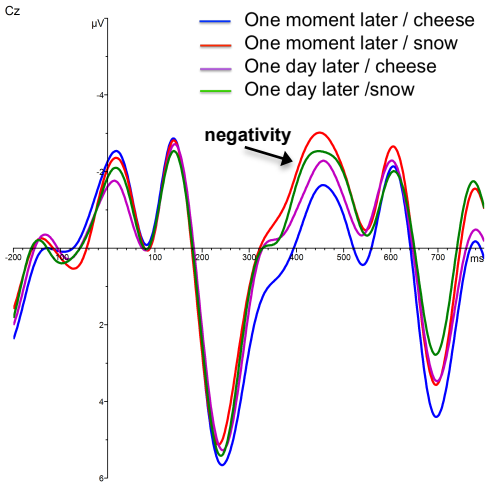


Figure 1: conditions 1-4 on electrode Cz; data was filtered for presentation purposes only with 8hz low-pass filter

Materials were normed in off-line experiments. Context-fitting and neutral objects were obtained in a cloze task study with high cloze values (>0.5) for script-fitting and low cloze values (<0.1) for neutral ones. A plausibility judgment task further ensured the acceptability of all target sentences.

Fitted linear mixed-effect models (300-500ms after onset of the direct object) show a significant, broadly distributed fronto-central negativity for conditions (2) and (4) compared to condition (1), as well as a frontal negativity (300-400ms) for major shift (3/4) vs. minor shift conditions (1/2). This is consistent with the view that script knowledge rapidly facilitates comprehension of script-congruent object nouns, which is not explainable by priming alone. Comprehenders appear to be sensitive to temporal components of scripts, making objects with an implausible temporal shift more difficult to process.

Results from Exp. 1 suggest that, even after a large temporal shift, a script-fitting object noun is still easier to process than a neutral one. One reason for this may be that the temporal shift used in Exp. 1 was not salient enough to completely deactivate a script. **Experiment 2**, for

which data is currently being collected, explores how script knowledge is used when context provides two scripts. One script is active, and thus expected to influence processing of target nouns to a greater extent.

Participants are presented with a context sentence introducing two scripts (Intro 1/2). By using a past tense structure combined with telic verbs, one script is perceived as ongoing, while the other is completed. A second context sentence provides more information regarding which point within either of the two scripts the target sentence is referring to.

The direct object of the word-by-word presented target sentence refers to an event either within the ongoing (1) or the inactive script (2). As suggested by reviewers, the 2x2 design is completed by the scripts in the context sentence presented in reversed order (Intro 2) to control for effects caused by the distance between script and direct object.

Intro 1: Script 2 active	**Intro 2: Script 1 active**
Nachdem Petra ins Kino gegangen war, erreichte sie die *Kneipe*.	Petra erreichte die *Kneipe*, nachdem sie ins Kino gegangen war.
(After going to the cinema, Petra reached the bar.)	(Petra reached the bar after going to the cinema.)
Rasch schaute sie auf den Aushang.	Rasch schaute sie auf den Aushang.
(Quickly she looked at the bulletin.)	(Quickly she looked at the bulletin.)

Target sentence:

(1)	Active-script target	Sie trat ein und kaufte das *Bier* ohne zu zögern. (She entered and bought the *beer* without hesitation.)
(2)	Inactive-script target	Sie trat ein und kaufte das *Ticket* ohne zu zögern. (She entered and bought the *ticket* without hesitation.)

Having shown that comprehenders are sensitive to temporal aspects of scripts (Exp 1), we expect to see a reduced negativity for the object fitting the active script (1) compared to the one fitting the inactive script (2). Any such effect would be unexplainable by simple priming, as both scripts are introduced.

We do not expect a main effect of script order, but merely an interaction with script congruency. This should manifest in a large N400 for inactive script-fitting objects when the active script was mentioned first (Intro 2 / Target 2) and a reduced N400 for script-fitting objects where the active script was mentioned last (Intro 1 / Target 1). The other conditions (Intro 1 / Target 2 and Intro 2 / Target 1) are expected to plot between these two extremes.

Since unexpected elements have been linked to positive-going ERP effects [8], we may additionally expect a positivity for the object fitting the inactive script.

By demonstrating that minimal linguistic material is sufficient to rapidly activate detailed script knowledge and make it accessible for language processing, we conclude that scripts provide an interesting method to investigate the interaction of non-linguistic knowledge in on-line comprehension. Specifically, drawing on aspects of their temporal and hierarchical structure we hope to further explore the role of implicit causal, temporal, and spatial relations in language comprehension.

References

[1] Kutas, 1993; [2] Kutas & Federmeier, 2011; [3] Schank & Abelson, 1975; [4] Rumelhart, 1980; [5] Bicknell et al., 2010; [6] Metusalem et al., 2012; [7] Kutas & Hillyard, 1984; [8] Thornhill & Van Petten, 2012;

Event participant relations in language and cognition: a cross-linguistic study

Before infants learn their first words, they represent events in terms of relations between event participants (Woodward, 1998). At the level of language, verbs highlight these participant relations: *love*, for example, relates someone who loves to something that is loved, a linkage often thought to be encoded in the verb's argument structure. In this study, we investigated the relationship between conceptual participant relations and verbal argument structure, focusing on the case of instruments, e.g. *Sam sliced the bread with a knife*. English instruments pattern like arguments given some argument diagnostics but pattern like modifiers given others. We asked whether this noncanonical argument status is reflected commonly across languages. Comparing the judgments of speakers of English, Spanish and Mandarin, we found strikingly similar results, supporting the theory that the interface between event concepts and verbal meaning is similarly constrained across languages, even for noncanonical participants such as instruments.

English instrumental verbs Instrumental phrases (e.g. *slice with a knife*) do not clearly pattern as either arguments or adjuncts (see Schutze, 1995). Koenig, Mauner & Bienvenue (2003), for example, propose that verbs such as *slice, beat* and *write* semantically "require" an instrument, in contrast with *eat, open* and *break,* which only "allow" an instrument. This require/allow-instrument contrast is not, however, reflected in syntactic argument behavior: the *with*-phrase patterns as an adjunct regardless of verb type (see Rissman, 2010).

Rissman, Rawlins & Landau (2015) investigated the semantic status of require-instrument verbs, testing the hypothesis that these verbs encode three arguments, parallel to dative verbs such as *lend* and *teach*. In a set of experiments, English speakers were instructed to report judgments about the number of "arguments" of a verb, e.g. that because *want* involves a person who wants and a thing wanted, *want* has two arguments. Subjects were directed towards a semantic rather than syntactic interpretation of this concept, e.g. they were told that "arguments" are crucial to the verb's meaning but are not necessarily mentioned in a sentence. Rissman et al. found that instruments were judged to be arguments more often for require- than allow-verbs, and that recipients were almost always judged as arguments of dative verbs. Nonetheless, there was a considerable gap in the argument judgment rates for require-instrument verbs vs. dative verbs: 35% vs. 87%. Given this contrast, Rissman et al. argue that instruments are not arguments, but are encoded gradiently as relatively salient extensions of the agentive force component of an event.

Spanish and Mandarin studies: method and results This proposal raises the question of whether gradient encoding of instruments is cross-linguistically common. That is, are there languages in which instruments are encoded as prototypical arguments, or where instruments are not argument-like at all? As a first step towards addressing this question, we adapted the materials from Rissman et al. (2015) into Spanish and Mandarin. We selected verbs in the same semantic space as the English instrumental and recipient verbs, and adapted the instruction about the nature of "arguments." Native speakers of Spanish and Mandarin (N = 28; 32) then read sentences such as in (1-4) and had to choose one of the bracketed phrases as an argument of the verb, or that neither phrase was an argument.

(1) Rachel REBANÓ algo [con una hoja de afeitar] [en el puerto].
Rachel slice-3PST something with a razor blade in the port

(2) 【在去年復活節那天】小琴 用【一把短柄小斧】⁞砍了⁞一些東西。
in last Easter Sunday Xiaoqin use one hatchet chop-PFV something

(3) [A las 6 am] Ruby le PRESTÓ algo [al nadador].
At 6 AM Ruby 3SG lend-3PST something to the swimmer.

(4) 克洛伊【在街上】⁞賣了⁞一樣東西【給演員】。
Chloe in street send-PFV something to actors

We asked two questions: 1) will all three languages show a strong asymmetry between recipients and instruments, and 2) will all three languages show gradient patterns across individual verbs? Given that the require- and allow-instrument categories were proposed on the basis of individual English verbs, we did not test whether comparable categories emerge in Spanish and Mandarin.

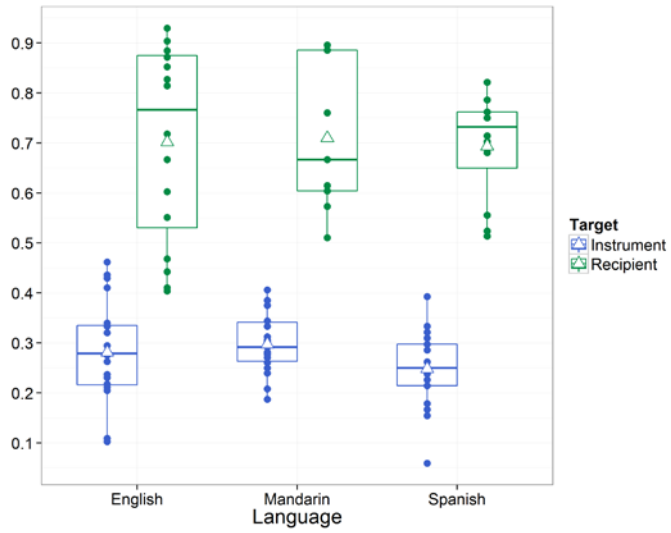


Figure 1. Rates of choosing the instrument or recipient as an argument in each language. Box plots show median and 2nd-3rd quartiles; triangles show means; dots represent single verbs

Figure 1 shows the rate of choosing the instrument or recipient as an argument for each verb and verb category, including the English data from Rissman et al. In a logistic regression model, we found that participants selected recipients more often than instruments as arguments ($\beta = 2.86$, SE = .10, $p < .001$), with no main effect of Language ($\chi^2(2) = .39$, $p > .1$). Thus neither Mandarin nor Spanish appears to encode the instrument as a prototypical argument. We also found significant variability within the verb categories in each language, suggesting instruments are generally encoded as more or less salient. Finally, we found significant correlations between the individual verb means across languages, suggesting that the semantic features of these verbs have comparable effects on event representation in each language (Recipients: English-Spanish = .84, English-Mandarin = .72; Instruments: English-Spanish =.81, English-Mandarin =.59; all p's < .01). Although our study only covers three languages, these findings support the theory that the relationship between conceptual event representation and verbal meaning is subject to cross-linguistically general biases. Encoding an instrument in an argument role may be dispreferred, raising questions about the conceptual foundations of such a dispreference.

References Koenig, Mauner, & Bienvenue (2003). Arguments for adjuncts. *Cognition.* | Rissman (2010). Instrumental *with*, locatum *with* and the argument/adjunct distinction. *LSA Meeting Extended Abstracts* 2010 | Rissman, Rawlins & Landau (2015). Using instruments to understand argument structure: evidence for gradient representation. *Cognition.* | Schutze (1995). PP Attachment and argumenthood. *MITWPL.* | Woodward (1998). Infants selectively encode the goal object of an actor's reach. *Cognition.*

Infants bind event aspects to nouns

The acquisition of verbs and nouns is considered to be differently difficult. While for nouns, it a mapping of lexical form onto the concept of the object is assumed, for verbs, different aspects of an event have to be recognized to become correctly linked with the lexical form. Accordingly, such views offer an advantage to the acquisition of nouns, since the mapping between a word and a concrete object (Gentner, 1982) results in a relatively stable meaning. In contrast, mapping of a verb to an event, is much more variable, and task-dependent. Recent developmental studies, have also provided new evidence in favor of this difference in acquisition (Bergelson & Swingley, 2012; Bergelson & Swingley, 2013).

Yet, this disadvantage in the acquisition of verbs stands in opposition to embodied cognition approaches positing the comprehension of verbal expressions being linked to mental simulation involving perceptual, motor and affective contents (Barsalou, 2008). Similarly, in developmental approaches, Mandler (2012) suggested that early concepts capture the roles of objects, i.e. what the objects do and what is done to them. In this way, actions are primary in organizing infants' concepts (Mandler, 2006). Even more recently, research has suggested that infants use relational properties instead of one-to-one mappings to learn new words (Yin & Csibra, 2015). This recent finding in a way echoes earlier, functionalist theories of language, according to which infants co-develop concepts about objects and events in dependence of the language input and the experiences they make in the world.

Following such an approach we reasoned that when learning new words, infants do not create simple correspondences with the external world but build relations within rich concepts. For our investigation, we hypothesized that (a) early event concepts are complex and comprise both actions as well as agents and objects involved in these actions and that (b) verbs are understood early in development but might be methodologically difficult to access.

Method

We replicated the Bergelson & Swingley (2012) study in which infants saw pictures of body parts and everyday objects and heard the mother's voice presenting verbal stimuli (see also Parisi & Cribra, 2012) with some modifications: Firstly, we used other pictures of everyday objects. Secondly, instead of nouns that are supposed to match the stimuli, in our study, the mothers provided verbs that can be associated with the presented objects. The verbs were contrasted along the dimensions CARE vs. ACTIVITY.

We tested 36 (18m;18f) 9-10-month-old infants (M = 9;25 ; SD = 19) using a looking while listening preferential looking paradigm with an eye-tracker (X2-60 Tobii). Infants saw two instantiations of paired pictures while hearing four verb pairs.

Results

We found that infants gazed at the target significantly more after hearing the target word then before ($t(35) = 2.26$; $p = .03$).

Our results show that 9-10-month-old infants are able to understand some verbs and refer them to object stimuli. This suggests that actions and objects are inherently linked: young infants can use contextual information (e.g. objects involved in an action) to infer the meaning of action words. Our study complements research

suggesting that children learn language by building relations and drawing from rich concepts.

References

Bergelson, E. & Swingley, D. (2012): At 6–9 months, human infants know the meanings of many common nouns. In: PNAS 109: 3253–3258.

Bergelson, E., & Swingley, D. (2013). The acquisition of abstract words by young infants. Cognition, 127(3), 391–397.

Mandler, J. M. (2006): Actions organize the infant's world. In: Hirsh-Pasek, K. & Golinkoff, R. M. (Eds.): Action meets word. How children learn verbs. Oxford: OUP: 111–133.

Mandler, J. M. (2012). On the Spatial Foundations of the Conceptual System and Its Enrichment. Cognitive Science, 36, 421–451.

Parise, E. & Csibra, G. (2012): Electrophysiological Evidence for the Understanding of Maternal Speech by 9-Month-Old Infants. Psychological Science 23: 128–733.

Werner, H. & Kaplan, B. (1963): Symbol formation: An organismic - developmental approach to language and the expression of thought. New York: Wiley.

Yin, J., & Csibra, G. (2015). Concept-Based Word Learning in Human Infants. Psychological Science, 26(8), 1316–1324.

EVENTS, ENTITIES, AND TEXT

Derry Tanti Wijaya (Carnegie Mellon University)
dwijaya@andrew.cmu.edu

When an event happens in the world, it involves a number of entities, some of which are then influenced or changed by the event (e.g., the person being affected by a shooting event). We have been studying how entities in the world change with respect to how they are described in text.

The series of studies that we conducted ask these questions: (1) can we track changes that happen to entities from the way they are described in text over time? (2) Can we find in the context within which a changed entity is mentioned in text, words that *trigger* the changes that happen to the entity? And are these trigger words indicative of the event that causes the changes? (3) Given the trigger words (we focus on verb phrases) and the changes in text that we learn to relate to a particular event, can we predict when the same event happens to other entity from its description in text? And can we automatically update our knowledge about the entity based on the changes brought about by the event?

We will present key findings from our studies with respect to each of the questions in turn. By presenting these series of studies, we are interested to start a discussion on how such verb- and entity-centric study of events in text can enrich the study of event structure in language.

Tracking changes that happen to entities in text

Tracking the frequency of an entity's mentions in a large collection of texts over time can indicate *when* a change is happening to the entity as its frequency changes [1]. Inspired by this earlier study, we seek to learn not only *when* but also *how* the entity changes from the change in the words that surround the entity over time [2]. We identify clusters of topics surrounding the entity over time and indicate that there is a topic change when two consecutive years are of different clusters.

We use K-means and Topics-over-Time clustering [3]; a Latent Dirichlet allocation (LDA)-style topic model that captures how latent structure in the data changes over time. For example, for the entity *Iran*, our clustering identifies a topic change in the year 1979 with the emergence of a new cluster consisting of words such as 'republic' and 'revolution'. Coincidentally, the year 1979 is the year of the revolution event that changed *Iran* from a monarchy to an Islamic republic. Another example, for the entity *Kennedy*, our clustering identifies a topic change in the year 1961 with a cluster containing words such as 'senator' transitioning to a new cluster containing words such as 'president'. Coincidentally, the year 1961 is the year *John F. Kennedy* was elected president. These findings suggest that we can identify *when* and *how* entities change over time in text and that these changes often coincide with events that happen to the entities.

We also observe that changes happen not only to entities in text but also to words that are adjectives or nouns such as *gay*. For the word *gay* our clustering identifies a topic change in the year 1970 with a cluster consisting of words such as 'happy' and 'lively' transitioning to a cluster with words such as 'lesbian', 'liberation', and 'movement'. Coincidentally, The year 1970 is also the year that the homosexuality movements began to pick up pace.

Finding words that trigger entities' change in text

In the following study [4], we seek to discover words that trigger entities' change in text and to find if these words are indicative of the event that causes the change. We start with some example entities; we call them *seed* entities that undergo a similar event at some points in in

their lives. We then aggregate the context of these seeds, which means averaging the frequency of words within which these 'seed' entities are mentioned in text: *before*, and *at* the time of the event.

Our hypothesis is that the context (surrounding words) *at* the time of the event contains words that are indicative of the event while the *difference* in contexts *before* and *at* the time of the event reflects changes associated with the event. Our observation confirms this hypothesis. For example, the aggregate context of seed entities for US presidency election contains words such as 'was elected', 'took office' which are indicative of the election event. We also observe that the *difference* in contexts before and after the event reflects the changes associated with the election event with the rise in mentions of 'administration' and 'president' surrounding the entity and the fall in mentions of 'senator', 'governor' and 'candidate' surrounding the entity. These findings suggest that from the frequency words surrounding the entities in text we can identify trigger words that are indicative of the event that causes the entities' changes.

Predicting change in entities

In this last study [5] we come full circle. We use trigger words that indicate a particular event to predict the change in the entities that is brought about by the event.

We focus on trigger words in the form of verb phrases that indicate events and the changes associated with the events. We learn these verbs and the associated changes from Wikipedia edit history. Our motivation is as follows: when an event happens to an entity, the *infobox* on the entity's Wikipedia page usually gets updated. At the same time, the article text may be updated with verbs either being added or deleted to reflect changes made to the infobox.

We use Wikipedia edit history to distantly supervise a Maximum Entropy classifier for automatically learning verbs and the associated changes that they bring. Our classifier learns that verb phrases such as 'die on' or 'pass on' trigger the deaths of their subjects (i.e., death dates are being added to their infoboxes). Verb phrases such as 'marry on' trigger their subjects to 'gain' a spouse (i.e., spousal information is being added to their infoboxes) while verb phrases such as 'divorce in' trigger the subjects to 'lose' a spouse. From the experiments, we observe that when these trigger verbs are being added or deleted from an entity's Wikipedia page, we can predict the entity's infobox changes with 88% precision and 76% recall. These findings suggest that we can automatically discover verb phrases in text that reflect events and that these verb phrases are effective for predicting entities' changes caused by the events.

References

[1] Michel, Jean-Baptiste, et al. "Quantitative analysis of culture using millions of digitized books." science 331.6014 (2011): 176-182.

[2] Wijaya, Derry Tanti, and Reyyan Yeniterzi. "Understanding semantic change of words over centuries." *DETecting and Exploiting Cultural diversiTy on the social web workshop*. ACM, 2011.

[3] Wang, Xuerui, and Andrew McCallum. "Topics over time: a non-Markov continuous-time model of topical trends." *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2006.

[4] Wijaya, Derry Tanti, Ndapandula Nakashole, and Tom M. Mitchell. "CTPs: Contextual Temporal Profiles for Time Scoping Facts using State Change Detection." *EMNLP*. Doha, Qatar. 2014.

[5] Wijaya, Derry Tanti, Ndapandula Nakashole, and Tom M. Mitchell. ""A Spousal Relation Begins with a Deletion of engage and Ends with an Addition of divorce": *EMNLP*. Lisbon, Portugal. 2015.

HOW EVENT ENDSTATES ARE CONCEPTUALIZED

Angela Xiaoxue He (Boston University) & Sudha Arunachalam (Boston University)

Many event descriptions are true only when the event comes to its natural end point. For example, an event of "feeding the baby" culminates only when the baby has eaten, and not simply when food is put on a spoon. A "throwing" event ends when the ball leaves the thrower's hand, while a "catching" event ends when it lands in the recipient's. How are the endstates of events conceptualized? In other words, when people view some event and construct a mental representation of it, how is the event endstate represented – as a critical component that defines the event category, or as a perceptual aspect, altering which doesn't change the category? In this paper, we report a study with English-speaking adults that suggests endstates are conceptualized as critical event components. We also delineate the design of an ongoing pilot study with English-learning infants that will reveal the origin of the adults' conceptualization. At the end, we discuss the insights this pursuit may provide on a bigger cross-linguistic picture.

In the current study with English-speaking adults, we ask: provided two events with the same ACTION but different ENDPOINTs, one ending with a naturally expected result (i.e. [+complete]), the other only partially achieved (i.e. [-complete]), do they perceive them as similar (and thus, likely to be members of the same event category) or different (and thus, likely to be different categories)? We used a Similarity Judgment Task, in which adult participants (n = 32) were shown pairs of scenes and asked to rate their similarity on a scale from 1 (most dissimilar) to 7 (most similar). In the *Experimental condition*, one scene depicted an action coming to its expected endpoint (e.g. a ball rolls into and completely knocks over a block tower), and the other depicted the same action but without achievement of the endstate (e.g. only some blocks are knocked over). In the *Control condition*, the ball rolled either behind or in front of the tower – a perceptually salient difference, but not related to endstate. *Filler trials* included two scenes that were identical, completely different (e.g., a girl dances vs. a boy waves), or had one salient change (e.g., a boy tosses a ball vs. an apple). The events were all presented in the absence of linguistic information.

We found that adults were more sensitive to changes relating to endstate than to perceptually salient differences unrelated to endstate, as reflected in a lower average rating score in the *Experimental* than in the *Control condition*; see Figure 1. These results suggest that different types of changes are registered differently in English-speaking adults' mental representations. Importantly, the difference between a complete and incomplete event is indeed registered, and carries more psychological weight than a mere perceptual difference.

To identify the developmental origin of such conceptualizations of event endstates, we have also designed an experiment for 14-month-old English-learning infants. We use a Habituation-Switch paradigm, in which we habituate infants to one event and switch it to the other to test for dishabituation. Infants are habituated to an event coming to its expected endpoint; in the *Experimental condition*, they are tested with an event with the same action but without achievement of the endstate, whereas in the *Control condition*, they are tested with an event with the same action and endstate but differing in some perceptually salient aspect. Dishabituation in the *Experimental condition* but not in the *Control condition* will suggest an adult-like conceptualization of event endstates.

This current investigation focuses on the English-speaking population. We should also be aware of the cross-linguistic differences in how event endstates are realized. In English, and some other languages (e.g., German), change-of-state predicates *entail* that the change of state

is complete; while in others (e.g., Hindi, Tamil), event completion is only *implicated*, thus allowing cancellation of the endstate (Arunachalam & Kothari, 2011; Ikegami, 1985; Pederson, 2007; Wittek, 2002). For example, in Hindi, it is possible to say 'Maya ate the apple' even if Maya only ate some of the apple, or to say 'Maya killed the fly but it isn't dead' (the fly is merely wounded), while these are notably odd in English. So, the semantic spaces of different languages are carved up in different ways: it seems English packages [+complete] and [-complete] into distinct lexical items, whereas Hindi merges them into a single lexical item. In the face of this cross-linguistic difference, a bigger research agenda emerges: in what ways are the conceptual spaces of speakers of different languages carved up; to what extent does the way the conceptual space is organized match the way the language's semantic space is organized; and how do these conceptual spaces develop? Our data in this study provides partial answer to this big question: the conceptual space of English-speaking is carved up in a way that places [+complete] and [-complete] events in distinct categories, matching up with the way English semantic space is carved up, and thus may provide a conceptual basis for English to package the ENDPOINT information into its predicates. Future work should also include speakers and learners of languages like Hindi to round out this investigation.

In sum, this study sheds light on the relation between language and cognition – in particular, how an understanding of event representation and non-linguistic cognition can inform us about the nature of linguistic representations, what the conceptual underpinnings are for the linguistic packaging of event components, and how these develop from infancy to adulthood.

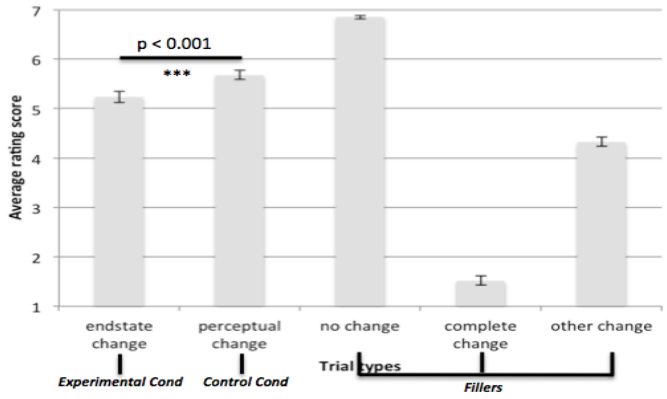


Figure 1
English-speaking adults' ratings of similarities between scenes

References

[1] Arunachalam, S. & Kothari, A. (2011). An experimental study of Hindi and English perfective interpretation. *Journal of South Asian Linguistics, 4*(1), 27-42. [2] Ikegami, Y. (1985). 'Activity'-'Accomplishment'-'Achievement' – A language that can't say 'I burned it but it did not burn' and one that can. In A. Makkai, A. K. Melby & R. Wells (Eds.), *Linguistics and philosophy essays in honor of Rulon S. Wells* (pp. 265-304). Amsterdam: John Benjamins. [3] Pederson, E. (2007). Event realization in Tamil. In P. Brown & M. Bowerman (Eds.), *Crosslinguistic perspectives on argument structure: Implications for learnability*. New York City, NY: Lawrence Erlbaum. [4] Wittek, A. (2002). *Learning the meaning of change-of-state verbs: A case study of German child language*. Berlin; New York: Mouton de Gruyter.

If You Want A Quick Kiss, Make It Count:
How Choice Of Syntactic Construction Affects Event Construal

When people talk to each other about what happened, they usually don't need to specify how long it took. Everybody knows from experience that a kiss lasts a few moments, a conference talk may carry on for about twenty minutes, and giving professional advice takes maybe half an hour, so there is typically no need to explicitly mention the duration. Duration is also usually not encoded grammatically. However, grammatical cues in event descriptions often significantly influence other aspects of event representations in listeners. It would be all the more interesting, thus, if very subtle grammatical choices were to reliably affect how long listeners think an event takes.

In this paper, we explore how encoding event descriptions in simple verbs (*to kiss, to advise*) versus count or mass noun light verb constructions (*to give a kiss, to give advice*) has repercussions on the temporal encoding of these events. Based on the fundamental observation that the reference properties of syntactic objects can change the reference properties of the whole predicate (Krifka, 1992; also see Quine, 1969; Verkuyl, 1972), we predict that nominalizing an event can help dividing experience into countable units, influencing duration estimates in a way that is systematically predictable from the interaction of verb semantics and nominal syntax.

A previous study has found that using count syntax but not mass syntax affects how events are quantified; and that punctive events are more readily quantified by counting over individual subevents than durative events. This is in line with the Number Asymmetry hypothesis (Barner and Snedeker, 2006): whereas count syntax specifies individuation, mass syntax is underspecified. We ask whether using a verb versus a noun in mass/count syntax affects how people judge event duration. The Number-Asymmetry hypothesis would predict that it would. However, if lexemes are ontologically linked to the same event structure, their syntax should not make a difference (Harley, 2003).

We used punctive (*hug*) and durative verbs (*advise*, *talk*) in either a transitive (*John hugged Mary*) or in a ditransitive frame with the bounded verb *give* (*give a hug*), which introduces a distinction between count (*give a hug/talk*) and mass syntax (*give advice*). We expected count syntax to force event individuation in punctives, such that, when asked about duration, people should judge the same event to be shorter in the ditransitive than in the transitive frame. For duratives, the same pattern was predicted for mass, but not count syntax: Since there are no distinctive subevents that can be counted, applying count syntax to duratives should not lead to differences in duration. Instead, it should open the door to different event construals, orthogonal to changes in temporal structure.

STUDIES 1 & 2: Duration Estimates. 100 English native speakers answered questions like *When they met up, Natasha hugged Cynthia. How long do you think that took?* Responses were transformed into log-seconds and analyzed with a mixed-effect model using maximal random effects structure. As predicted, punctive count syntax and durative mass syntax (βs>.34, ts>1.8, ps<.05), but not durative count syntax (β>.19, ts>1.1, ps>.25), led to shorter event construal (Figure A). Study 2 was a replication with verb type as between-subjects factor, yielding the same pattern of results.

STUDY 3: Duration Categorizations. To ensure that the pattern obtained by free duration estimates was not due to unreliable time estimates (Kruger & Evans, 2004), we provided 100 participants with pre-defined time bins as choices. The choices were each transitive/ditransitive pair's individual quartiles obtained in Study 1 as possible answers, e.g.: *When they met up, Natasha hugged Cynthia. How long did this take? a) up to 10 seconds, b) between 10 seconds and 30 seconds, c) between 30 seconds and 10 minutes, d) more than 10 minutes*. For punctive count pairs, there was a significant difference in categorizations depending on construction. For durative count and durative mass items, this difference was not significant; however, the effect of construction went in opposite directions for durative count items (β=-.29, p>.26), compared to punctive count (β=0.81, p<.03) and durative mass items (β=0.61, p>.14): whereas the ditransitive construction resulted in more "shortest" categorizations for punctive count and durative mass items, this effect was absent (numerically: reversed) for the durative count items (Figure B).

STUDIES 4 & 5: Event Repetition. 80 participants read each sentence and then noted how many events they imagined. Mean count of events was lower for ditransitive light verb constructions in each event category. For punctive count events, using a ditransitive light verb construction instead of the transitive verb reduced the mean count from 2.3 to 1.8 (χ^2=7.18, p<0.007). These effects were also numerically present for durative count and durative mass items, but variances in these constructions were higher (χ^2s<1.1, ps>0.29, Figure C). Study 5 was a replication asking for the specific event in question (e.g., how many kisses), yielding similar results.

STUDY 6: Event Similarity. We asked 100 English native speakers to rate event similarity between transitive and ditransitive frames on a 7-point Likert scale (1="same event"). As predicted, differences between frames were smaller in punctive verbs with count syntax (*to hug* vs *to give a hug*, mean rating: 1.55) and durative verbs with mass syntax (*to advise* vs *to give advice*; mean rating: 1.56) than in durative verbs with count syntax (*to talk* vs *to give a talk*; mean rating: 2.17; βs>.6, ts>4.8, ps<.0001; mixed-effect model with maximal random effects structure; Figure D).

SUMMARY: Describing an event with mass/count syntax affects event construal in a way that is systematically predictable from the interaction of mass/count syntax and semantics: *give a hug/advice* are imagined as taking less time than *hug/advise*; this does not apply to (*give a) talk*, supporting the Number-Asymmetry hypothesis, but not a hypothesis in which lexemes have the same ontology, independent of syntax, and complementing studies suggesting that people conceptualize events differently depending on subtle choices among alternations (Johnson & Goldberg, 2012; Wittenberg & Snedeker, 2014).