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

Eva Wittenberg* & Roger Levy

Department of Linguistics
University of California, San Diego
9500 Gilman Drive
La Jolla, CA 92093-0108

*Corresponding Author; electronic address: ewittenberg@ucsd.edu.

Abstract

When we hear an event description, our construal of the event is not only based on the lexical items in the message, but also on the message's syntactic structure. This has been well-studied in the domains of how listeners conceptualize causation, roles of event participants, or object properties, depending on syntactic choices. Less studied are subtle aspects of the construal of temporality and numerosity of events as a function of syntax. This paper presents a theory of how the syntax of a event description affects the construal of event similarity and duration in a way that is systematically predictable from the interaction of mass versus count syntax and verb semantics, and tests these predictions in six studies. We found that punctive events in count syntax (*give a hug*) and durative events in mass syntax (*give advice*) are construed as taking less time than in transitive verb frame (*hug* and *advise*). Durative verbs in count syntax, on the other hand, yield a different pattern: Count syntax (*give a talk*), compared to simple transitive constructions (*talk*), leads to a semantic shift in event structure and does not systematically affect duration estimates. These results support our theory of how syntactic and semantic structure together affect event construal.

Keywords Events, Aspect, Light Verb Constructions, Mass-Count distinction, Individuation, Linguistic Framing

1 Introduction

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, such as causation, event structure, and the semantic roles of event participants (Fausey and Boroditsky, 2010; Johnson and Goldberg, 2013; Wittenberg and Snedeker, 2014). 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.

This hypothesis was inspired by a previous study on how events are individuated, depending on mass and count syntax. Barner, Wagner, and Snedeker (2008) found that using count syntax *(to do climbs)*, but not mass syntax *(to do climbing)*, affects how events are quantified; and that punctive events (*kissing, kicking*) are more readily quantified by counting over individual subevents (*kisses, kicks*) 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.

If it is true that mass and count syntax contribute to event individuation, then we should expect predictable influences of mass or count syntax also on estimates of event duration. Figure 1 provides a schema of our predictions. We distinguish between two types of events: Shorter punctive, mostly semelfactive events, like kissing or kicking; and longer durative, mostly activity events, like talking or advising (Dowty, 1991; Vendler, 1957).

Punctive events are distinct from durative events not only in that they

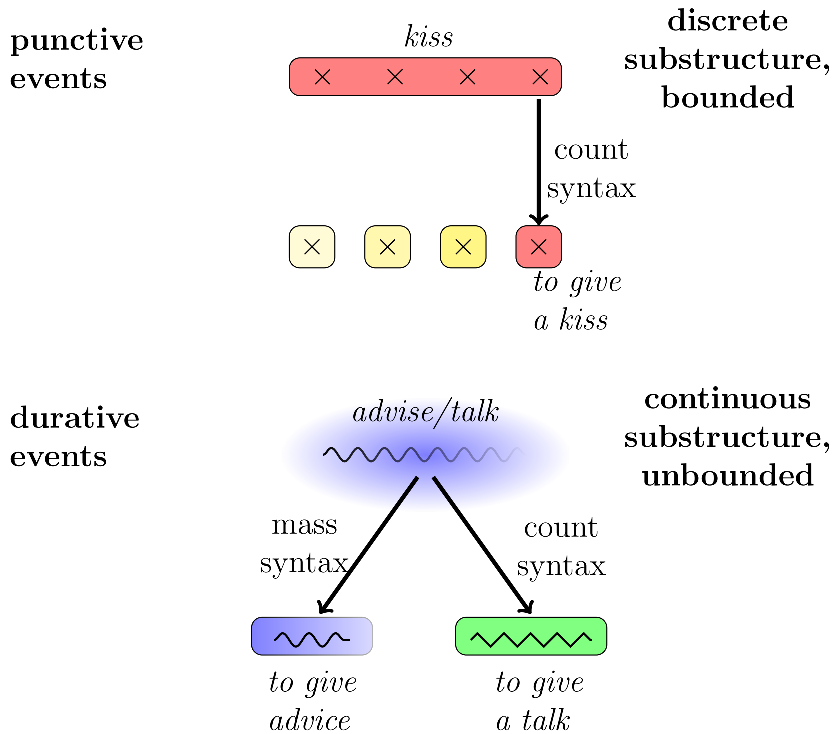

Figure 1: Predictions of how mass versus count syntax interact with verb semantics, with regards to event duration construals. See Section 1.3.

are short and bounded by a natural end point, but also in that sentences in which they appear are often conventionally understood to describe several instances of the same punctive event, that is, they are understood iteratively (Barner, Wagner, and Snedeker, 2008; Kim and Kaiser, 2015; Paczynski, Jackendoff, and Kuperberg, 2014): For instance, you may find that *Douglas kissed Mary* evokes the image of not one, but multiple kisses, each of which can be categorized as a subevent of kissing. Thus, punctive events can have a distinct substructure. Durative events, in contrast, are unbounded and continuous over time, they do not possess a distinct substructure, and they do not receive an iterative reading, even if the duration of the event is explicitly extended beyond a conventionally accepted time frame (cf. *Senator Cruz talked all night*).

Many of these events, like *kiss, advise* or *talk*, can either be expressed as transitive verbs, or as so-called light verb constructions. In light verb constructions, the verb contributes little semantics beyond tense, number agreement, and aspect, while the meaning of the expression comes from the deverbal noun (Brugman, 2001; Butt, 2003, 2010; Jackendoff, 1974; Jespersen, 1954; Wiese, 2006). These light verb constructions appear either with count syntax, such as *to give a kiss* and *to give a talk*, or mass syntax, such as *to give advice*. Thus, light verb constructions offer us an opportunity to study the interaction of verb type and mass versus count syntax with an existing alternation, as opposed to unusual constructions such as *to do climbs* (Barner, Wagner, and Snedeker, 2008): Light verb constructions, like *to give a kiss*, and their full verb counterparts, like *to kiss*, are in a relationship of syntactic alternation with minimal difference in meaning (Allerton, 2002; Glatz, 2006). In our study of punctive and durative events, we use light verb constructions with *give*, which is itself bounded (Newman, 1996).

1.1 The mass-count distinction and verbal aspect

As has been noted in the literature, there are interesting parallels between the boundedness or unboundedness of events in time, and count and mass noun phrases (Hale and Keyser, 1993; Harley, 2005; Jackendoff, 1991; Krifka, 1992): Both count nouns and bounded events have relatively precise limits and set borders: *an apple* and *a talk* are analogous to *fall asleep* and *sneeze* in that their limits are well defined in space and time, respectively. For mass nouns and unbounded events, this is not the case: *wine* and *advice*, for example, have no set borders, and neither do unbounded verbs, like *run*.[1]

Given these parallels, and given the bounded verb *give*, we expect count and mass syntax to interact differently with punctive and durative event nouns: Count syntax with a punctive deverbal noun like *giving a kiss* should pick out one single instance of *kissing*; mass syntax with a durative deverbal noun like *giving advice* should carve out a portion of *advising*. If this is true, then the grammar conveys a subtle difference in event duration between the simple transitive verb and the light verb construction: the light verb construction communicates an event that lasts a shorter time.

For durative verbs that have count nominals with the bounded light verb *give*, however, the predictions are not as straightforward. Just consider what

[1]The boundedness of mass nouns is introduced when they are quantized (*a bottle of wine, a piece of advice)*; see Krifka (1992); Wiese and Maling (2005).

happens when one packages substances like *beer, glass, string, stone* or *iron* into count syntax: In some cases, the count noun denotes portions (of arbitrary size) of the substance (*a string, a stone, a beer*), but in other cases, the count noun phrase happens to encode or units that are related, yet arbitrary, such as in *a glass* or *an iron.* Crucially, the resulting denotation for mass nouns in count syntax is variable, and each case conveys aspects of meaning that cannot be predicted from the intrinsic structure of the underlying substance (Gordon, 1985; Srinivasan and Rabagliati, 2015).

So what should we expect in the case of durative verbs entering count syntax, like *talk* in *give a talk*? If the analogy between durative verbs and substance nouns really holds, one would predict a certain degree of conventionalization, similar to the *some glass* → *a glass* case: Core parts of a given durative event would remain the same, but the event type would shift in meaning from verb to count noun construction, possibly closing a lexical gap in doing so (Allerton, 2002; Glatz, 2006; Grimshaw and Mester, 1988; Miyagawa, 1989). For example, there is a strong intuition that *giving a talk*, albeit still retaining the core meaning of utterance production, is conceptually further from *talking* than *giving a kiss* is from *kissing* (see Experiment 4).

But then, if it is true that durative events in count syntax undergo a conceptual shift in event kind, predictions about event duration are up in the air, since the change induced by count syntax would be orthogonal to changes in temporal conceptualization. In this paper we explore both sides: Whether event duration estimates are modulated by the introduction of mass and count syntax, and whether there is an effect on how similar events are judged as depending on the syntactic construction. In the next section, we will discuss the link between event representation and linguistic encoding in more detail.

1.2 Event construal via linguistic encoding

How linguistic framing influences people's event conceptualization, memory, and recall has long been a topic of interest in science, such as in behavioral economics (Halkjelsvik, Jørgensen, and Teigen, 2011; Kahneman and Tversky, 1977; Kruger and Evans, 2004; Roy, Christenfeld, and McKenzie, 2005), criminal justice (Boltz, 1995; Loftus et al., 1987; Macar, Grondin, and Casini, 1994; Tse et al., 2004), and cognitive science (Madden and Zwaan, 2003; Magliano and Schleich, 2000). These studies, however, were mainly concerned with behavioral or memory consequences of linguistic encoding accompanying visual scenes, and less with representational or grammatical issues.

In psycholinguistics, using grammatical alternations to study their influence on event construal started with Gropen et al. (1991), who found that subtle changes to event structure affected which form of the locative alternation people used in production (*cover a surface with marbles* or *dropping marbles onto a surface*). A later study confirmed the intuition that syntactically omitting agents from an event description reduces how much blame is assigned to them (Fausey and Boroditsky, 2010). Finally, directly related to the constructions used in this paper, we know that using a light verb construction with *give* influences the construal of thematic roles (Wittenberg, Khan, and Snedeker, under review; Wittenberg and Snedeker, 2014).

The question of whether differences in event descriptions cause differences in duration estimates is less well studied. So far, there is only a small but growing experimental literature showing that verbal aspect or the verb itself influence duration estimates. For example, human locomotion events seen on video are remembered as taking longer when they are described as walking events than when they are described as running events (Burt and Popple, 1996); and a Dutch study has shown that describing a short event in progressive verb aspect *(is kissing)* makes people think that the kiss took longer than if simple present is used (Flecken and Gerwien, 2013). Pedersen and Wright (2002), in contrast, found only small effects of event descriptions on the duration estimates; however, their manipulation was on writing style and purposefully not as tightly controlled for semantic and syntactic factors as other studies.

Thus, there is some evidence that linguistic choices influence the way people think about the temporal dimensions of events. Yet, it is not surprising that one should find this influence by grammatical means that by default operate in the temporal dimension, like aspect, or by choosing lexical items according to the speed of an action that they express. The alternation between transitive verbs and their light verb construction counterparts, however, affords a way to look beyond the more obvious aspects of how syntactic and semantic regularities work together to create a rich, full, and detailed representation of an event.

How could this interaction work? Most current theoretical approaches have accounted for the fact that some lexical items can appear both as nouns and verbs with the stipulation that prior to lexical insertion, their grammatical status is neutral. That means, a word like *kiss* only acquires its grammatical category upon insertion into a syntactic tree; either way, regardless of whether it is inserted as a noun (such as *He gave her a kiss*) or a verb (*He kissed her*), it conceptually refers to the same kind of event (Barner and Bale, 2002; Halle and Marantz, 1994).

Taking this as a starting point, we should be able to observe a systematic interaction between syntax, the lexicon, and event construal.[2] We assume that syntactic, conceptual, lexical, and phonological structure interact and predict upcoming features and structures every step of the way (Garrod and Pickering, 2003; Jackendoff, 2002, 2007; Levy, 2008). This architecture of the linguistic system would allow for an interaction between representations on the levels of syntax (mass and count syntax), semantics (the boundedness of *give*), and event knowledge (how long kissing usually takes). Other approaches (Goldberg, 1995, 2003) explicitly allow for different meaning shades of grammatical constructions; presumably, the distinct meaning of a ditransitive construction as a bounded event would be even more straightforwardly predicted from this perspective. Crucially, both of these models would predict differences in duration estimates due to the interaction of verbal and nominal semantics and syntax.

1.3 Key theoretical predictions and the current studies

The current studies investigate whether describing an event with mass or count syntax, as opposed to a simple transitive verb, affect the construal of event duration, event similarity, and event repetitions in a comprehender. Figure 1 shows a graphical representation of our theory.

Starting from the top of Figure 1, we look at punctive events like *kissing*. Previous studies have shown that punctive events are often construed as occurring more than once (Barner, Wagner, and Snedeker, 2008; Kim and

[2]Another option would be a strictly linear model: A lexical item is inserted into a syntactic tree, its phonological form created, and its meaning read off by the comprehender, without any feedback loop to the semantic or conceptual system. In this case, both forms of the lexical items (e.g., *kiss*, *talk* or *advise* as either noun or verb) are ontologically linked to the same event structure regardless of grammatical status, and it should not make a difference for duration estimates whether they appear as a verb or in either mass or count noun syntax, since the comprehender's lexicon is set up the same way – *kiss* conveys the same event regardless of whether it appears as verb or noun. Thus, this linear model would not predict any differences in duration estimates.

Kaiser, 2015; Paczynski, Jackendoff, and Kuperberg, 2014), e.g., a comprehender might hear *Douglas kissed Mary* and imagine more than one kiss. In Figure 1, this is represented by four crosses, which are distinctive subevents within one bounded kissing events. Count syntax, according to our theory, should encourage event individuation in iterative events: In the case of our example, *giving a kiss* describes only one particular kiss. This in turn should have repercussions in the construal of event temporality, leading to conceptualizations of shorter event duration. Another prediction is that punctive events in count syntax will be construed as consisting of fewer event repetitions than in transitive syntax.

The bottom of Figure 1 visualizes our predictions for durative events. Durative events, represented by the continuous snake line, have no set endpoint (and in many cases, no set starting point either). They also do not have a continuous substructure; for example, *advising* is a process that consists of many points in time, but it is hard to conceptually delimit when advising starts and ends based on these time points. However, it is possible to carve out a particular portion of this process, and we claim that this is linguistically done when the event appears in mass syntax: *to give advice* refers to a chunk of advising whose boundaries limited (although not quite as strictly as in the case of *to give a kiss*). In terms of event counts, the predictions are weaker than for the punctive event counts, since mass syntax does not aid in event individuation (although the boundedness of the verb *give* might).

When durative events occur in count syntax, such as in *to give a talk* (bottom right on Figure 1), we hypothesize that, analogous to the cases of using mass nouns in count syntax (*glass, iron* versus <u>*a*</u> *glass,* <u>*an*</u> *iron*), there are arbitrary and unpredictable changes in the kinds of events these nouns then describe. For example, *giving a talk* still retains a sense of utterance production, but in a very different context and with different event parameters (this difference is represented by a change in color, and the zigzag line, in Figure 1). This creates another prediction: *Talking* and *giving a talk* should be conceptually less closely related than punctive count and durative mass pairs. In terms of event counts, we do predict that the count noun form will help event individuation and possibly lead to a reduction in event counts. Crucially, both event similarity and event count differences would operate entirely orthogonally from the construal of event duration.

We present six experiments that test these predictions. To test the claim that the construal of event duration is predictable from the interaction of mass versus count syntax, and verb semantics, we present two studies that elicited open estimates of event duration (Section 2, Experiment 1a and 1b); then, we present a categorization experiment, in which participants categorized event descriptions into predefined time bins (Section 3, Experiment 2). Section 4 presents two studies that establish whether using count or mass syntax affects how many events people imagine (Experiment 3a and 3b). Finally, in Section 5, we investigate whether describing durative events in count syntax indeed leads to conceptual shifts in event type (Experiment 4).

1.4 Stimuli used for all studies

	count	**mass**
punctive	*to kiss – to give a kiss*	
	to embrace – to give an embrace	
	to hug – to give a hug	
	to kick – to give a kick	
	to poke – to give a poke	
	to shake – to give a shake	
	to cuddle – to give a cuddle	
durative	*to talk – to give a talk*	*to advise – to give advice*
	to address – to give an address	*to thank – to give thanks*
	to lecture – to give a lecture	*to assure – to give assurance*
	to present – to give a presentation	*to encourage – to give encouragement*
	to speak – to give a speech	*to recognize – to give recognition*
	to scold – to give a scolding	*to support – to give support*
		to assist – to give assistance

Table 1: Experimental item pairs used for all experiments.

We used punctive verbs (*kiss*) and durative verbs (*advise, talk*) either in a transitive frame (*After their first date, Douglas kissed Mary*) or in a ditransitive light verb construction with the bounded verb *give* (*After their first date, Douglas gave a kiss to Mary*) such as in Table 1 (see Appendix A for a full list of stimuli). The ditransitive frame introduces a distinction between count syntax (*give a kiss/talk*) and mass syntax (*give advice*). We expected count syntax to force event individuation in punctive verbs, such that, when asked about event duration, people should judge the same event to be shorter in the ditransitive than in the transitive frame. For durative verbs, we predicted the same pattern for mass syntax, albeit to a lesser degree, because only the light verb *give* would encourage boundedness. For durative events that enter count syntax, we predicted a different pattern: Since there are no distinctive subevents that can be counted, applying count syntax to durative verbs should not lead to differences in duration. Instead, it should open the door to different event construals, orthogonal to changes in temporal structure (see Figure 1 for an overview of predictions).

Of the twenty experimental items, seven encoded punctive events, e.g. *kissing/giving a kiss*, six were durative events and could be used with count syntax, e.g. *talking/giving a talk*, and seven encoded durative events that could be used with mass syntax, e.g. *advising/giving advice*.[3] Note that in this last category, the noun describing the action (*advice*) was not preceded by a determiner. None of the light verb constructions that we used can alternate between mass and count syntax (e.g. *to give *(a) kiss; to give (*an) advice*). In short, there were three different verb alternations: punctive transitive vs. ditransitive with count syntax (PC); durative transitive vs. ditransitive count syntax (DC); and durative transitive vs. ditransitive mass syntax (DM).

In all studies except for Experiment 4, the items were embedded in a sentence context. All sentences used the simple past. Each sentence included a temporal or local adjunct phrase, to encourage non-repetitive readings (e.g. *The professor advised her student on his paper yesterday*). In addition to the experimental items, we also created 27 filler sentences.

2 Experiment 1a: Open estimates of event duration

In Experiment 1a, we asked participants to rate how long events took. The events were described by simple transitive or light ditransitive constructions, and we were interested in whether the predictions displayed in Figure 1 would be confirmed. If so, we expect punctive events like kissing to be estimated as taking less time in the ditransitive construction (*giving a kiss*). We should also observe the same trend for durative events when they occur in mass

[3]The missing item in the durative count category was *condole – giving condolences*, which we ultimately decided to exclude because of its markedness as a verb. In English, there are no punctive events in mass syntax with *give*.

syntax (*advising* – *giving advice*), but not durative events in count syntaxx (*talking* – *giving a talk*). Experiment 1b served as a replication, with event category as a between-subjects factor.

2.1 Methods

2.1.1 Participants

We recruited 100 unique individuals on Amazon Mechanical Turk, an online crowd-sourcing tool. Mechanical Turk allows access to a large number of study participants, who participate anonymously for reasonable compensation, in our case for about $6 an hour (Buhrmester, Kwang, and Gosling, 2011; Crump, McDonnell, and Gureckis, 2013). Our participants had IP addresses within the United States and were self-reported native speakers of English.

2.1.2 Stimuli

We created two lists out of the sentences described in Section 1.4 and distributed the experimental items across them with a Latin-square design, such that each participant saw every sentence in either transitive or ditransitive light verb form. The fillers were the same across lists.

2.1.3 Procedure

Participants read each sentence and then estimated how long the event described in the sentence probably took. The exact instructions were:

> "In this survey, we would like your opinion on how long you think events take. For example, the event "Frida is eating an apple" might take about seven minutes; "Frida is driving 150 miles" might take about 2 hours 15 minutes; or "Frida sneezed" might take about one second. You can write your estimates into the hour, minutes, and seconds boxes for each event, so please read each of the following sentences carefully and estimate how long the described event took as best and accurately as you can."

For each item, participants were able to enter their estimated event duration in a set of three text boxes, one for hours, one for minutes, and one for seconds, e.g., a participant could respond "1 hour(s), 13 minute(s), 7 second(s)". Empty boxes were treated as a response of zero for that unit of time. Completing the study took about 17 minutes on average.

2.2 Results

Responses in which the estimated duration was zero were discarded. This affected less than .1% of the data. Since effects of grammatical structure on event duration would be likely to operate proportionally to intrinsic event duration, we transformed all responses to log-seconds for purposes of data summarization and analysis. Figure 2 shows the pattern of results, with responses back-transformed to hours, minutes, and seconds, for convenience of interpretation.

For punctive count events, using a ditransitive light verb construction instead of the transitive verb cut the time estimates in half, from about 40 to about 20 seconds. The difference was almost as pronounced for durative mass items (transitive μ=50 minutes, ditransitive μ=29 minutes), but for durative count items, the effect was absent (transitive μ=31 minutes, ditransitive μ=27 minutes).

For statistical analysis, we used linear mixed-effects regression models (Baayen, Davidson, and Bates, 2008) with maximal random effects structure, including event category (PC, DC, and DM), construction (transitive or ditransitive light verb) and their interaction as fixed effects, random by-participant intercepts and slopes for all fixed effects, and random by-item intercepts and slopes for event category (Barr et al., 2013), using the `lme4` package (Bates et al., 2014). We used Helmert coding for both fixed-effects predictors. For event category, we grouped PC and DM items together as one Helmert contrast pair, and their average contrasted with DC items as the second Helmert contrast pair.

The top half of Table 2 shows the results of likelihood-ratio tests for the two main effects and the interaction, comparing the full model specified above with reduced models in which either main effect, or the interaction, was removed from the fixed-effects structure.[4] For these three tests, we removed all random correlation parameters from both null- and alternative-hypothesis models to ensure model convergence.

The bottom half of Table 2 shows the planned pairwise comparison results

[4] See Levy (2014), among others, for details on how omnibus main effects can be tested in the presence of an interaction in `lme4`.

for each event category. The results show that whenever punctive events appear in a count light verb construction with *give* (ditransitive frame, such as *to give a kiss*), they are estimated to take less time than when they appear in a simple transitive (*to kiss*). The same is true for durative events in mass syntax (*to give advice – to advise*), but not for durative events in count syntax (*to give a talk – to talk*).

	Df	χ^2	p-value
construction	1	7.9	0.005 ***
event category	2	23.02	0.000 ***
construction × event category	2	3.28	0.196 *n.s.*
PC – construction	1	3.54	0.059 .
DC – construction	1	0.15	0.69 *n.s.*
DM – construction	1	6.17	0.013 *

Table 2: Table of likelihood estimation results for duration estimates in Experiment 1a, testing the main effects and their interaction (upper part), and the results of testing the main effect of construction in pairwise comparisons (lower part).

2.3 Replication: Experiment 1b

We replicated this study with 300 participants on Amazon Mechanical Turk by making event type (PC, DC or DM) a between-subjects factor, reasoning that thinking about events on vastly different time scales might wash out sharper judgments. The exact instructions were:

> "In this survey, we would like your opinion on how long you think events take.
>
> Please take a look at your computer clock or watch, and think about how long thirty minutes take. What can you do in thirty minutes? You might be able go grocery shopping, for example. But the event of grocery shopping might take much shorter (if you only need a few items and don't have to wait at the cashier, you might be in and out in ten minutes), or much longer (if you need a lot of things and there is a long line, you can easily spend two hours).

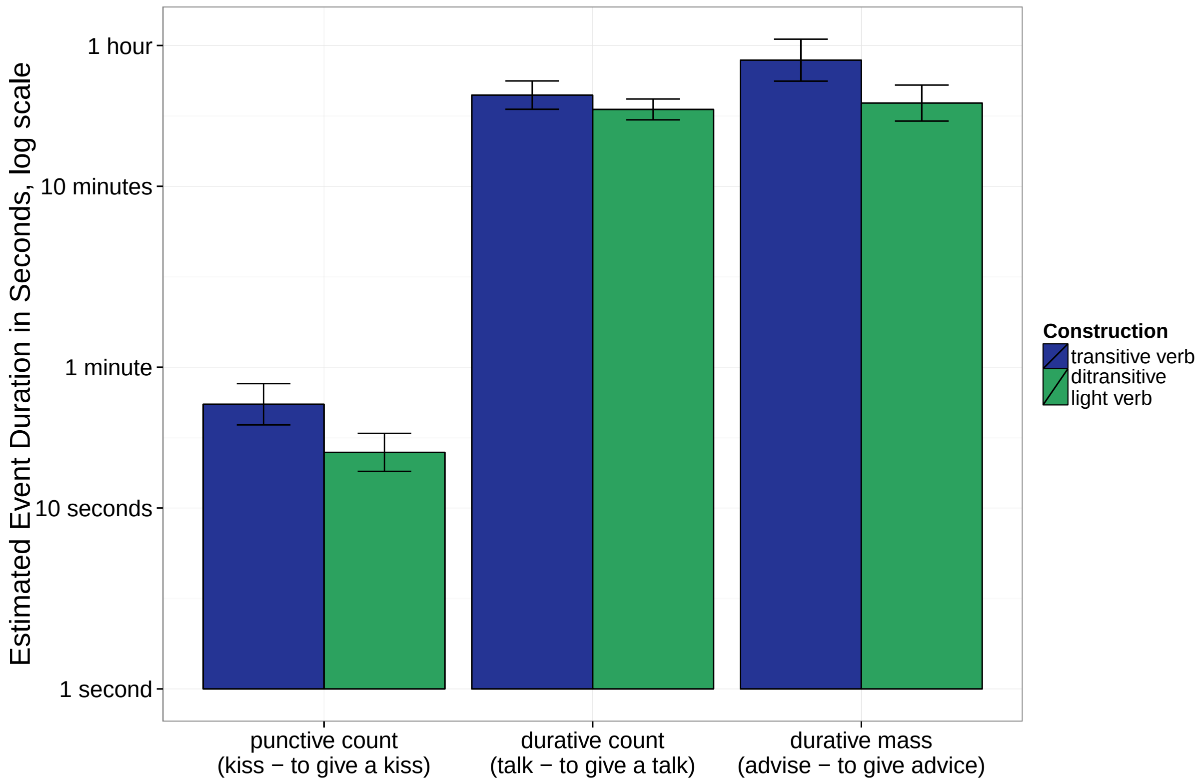

Figure 2: Experiment 1: PC and DM item pairs were rated to take less time when they were presented in ditransitive light verb frames. For DC item pairs, the syntactic frame made no difference. The y-axis is represented in log scale, but labeled with human-readable time estimates for convenience. Error bars represent 95% confidence interval.

> In this experiment, we would like you to think how long events *typically* take. You can write your estimates into the hour, minutes, and seconds boxes for each event, so please read each of the following sentences carefully and estimate how long the described event took as best and accurately as you can."

Figure 3 shows the pattern of results; again, punctive count item pairs were estimated to take roughly half as long when they were presented in ditransitive light verb frames (transitive μ=14 seconds, ditransitive μ=27 seconds). For DC and PC item pairs, the syntactic frame made no statis-

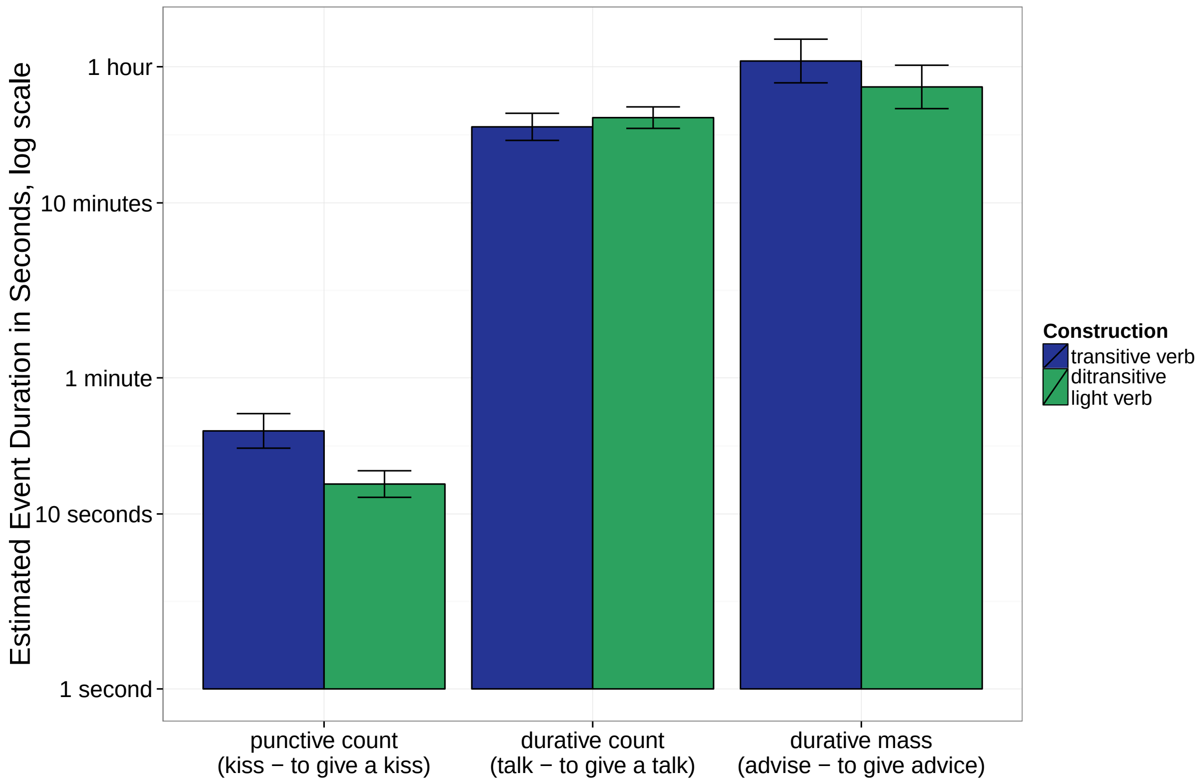

Figure 3: Experiment 1b (replication with event category as between-subjects factor): Duration Estimates in log-seconds, per item pair. Error bars represent 95% confidence interval.

tically significant difference (see analysis below): Durative mass items were estimated to take 76 minutes in the transitive frame, and 54 minutes in the ditransitive frame, and durative count items were estimated to take 40 minutes in the transitive frame, and 27 minutes in the ditransitive frame.

Table 3 shows the results of the regression analyses. Like in Experiment 1a, we separately tested for main effects of construction and event category in a linear mixed-effect model, including event category (PC, DC, and DM), construction (transitive or ditransitive light verb) and their interaction as fixed effects, random by-participant intercepts and slopes for all fixed effects, and random by-item intercepts and slopes for event category, using the `lme4` package. We used Helmert coding for both fixed-effects predictors. Again, we grouped PC and DM events together as one Helmert contrast pair, and their average contrasted with DC items as the second Helmert contrast pair. Event category was a between-subjects factor.

The top half of Table 3 shows the results of likelihood-ratio tests for the two main effects and the interaction, comparing the full model specified above with three reduced models in which either main effect, or the interaction, were absent from the fixed-effects structure. Again, we removed all random correlation parameters from both null- and alternative-hypothesis models to ensure model convergence. Both main effects of construction and event category, and their interaction, were significant.

The planned pairwise comparison results are shown in the bottom half of Table 3. Again, whenever punctive events appear in a count light verb construction with *give* (ditransitive frame, such as *to give a kiss*), they are estimated to take less time than when they appear in a simple transitive (*to kiss*). The effect was not significant for durative events with mass syntax (*to give advice – to advise*), or durative events with count syntax (*to give a talk – to talk*).

	Df	χ^2	p-value
construction	1	4.76	0.029 *
event category	2	27.58	0.000 ***
construction $\times$ event category	2	6.00	0.049 *
PC – construction	1	9.64	0.001 **
DC – construction	1	0.40	0.531 *n.s.*
DM – construction	1	0.95	0.330 *n.s.*

Table 3: Table of likelihood estimation results for duration estimates in Experiment 1b, testing the main effects and their interaction (upper part), and the results of testing the main effect of construction in pairwise comparisons (lower part).

2.4 Discussion of Experiments 1a and 1b

As predicted, punctive events are estimated to take longer when they appear in transitive syntax than when they appear in a count noun frame (*kissing* vs. *giving a kiss*). The same trend was observed for durative events when they appear in mass syntax: *advising* is estimated to take less time than *giving* *advice*. Duration estimates for durative events are not affected, though, when these events appear in count syntax (*talking* vs. *giving a talk*). We were able to replicate the same pattern of results with event type (PC, DC, and DM) as between-subjects factor, although the effect did not reach significance for DM items.

One might notice that the duration estimates seemed fairly high. We know from previous studies that estimating the duration of an event is often influenced by its pleasantness or desirability (Kahneman and Tversky, 1977; Roy, Christenfeld, and McKenzie, 2005). Although the numerical values of the estimated event durations are not crucial here, since we are only interested in the difference in estimates that is attributable to syntactic structure, it is conceivable that our participants' estimates were influenced by factors not due to the grammatical construction alone. For this reason, we decided to replicate this study as a categorization task with predefined time windows as answer options. The predefined time windows were based on previous answers, to validate the results obtained in Experiment 1a and 1b.

3 Experiment 2: Categorizations of event duration

This experiment gave participants the opportunity to estimate event durations without needing to come up with their own time estimates, and instead being able to select among predefined time bins for each individual event.

3.1 Method

3.1.1 Participants

We recruited 80 self-reported native speakers of English on Amazon Mechanical Turk with IP addresses within the United States.

3.1.2 Stimuli

We used the same sentences as in the previous studies. Participants read each sentence and then had to categorize the event for duration by clicking one of four options. These options were created using the quartiles of each item pair's averaged estimated durations in Experiment 1; thus, every item had different answer options, as in Examples (1) and (2):

(1) *Laughing nastily, the thug kicked the victim.*
How long did this take?

a) up to 5 seconds
b) between 5 seconds and 13 seconds
c) between 13 seconds and 10 minutes
d) more than 10 minutes

(2) *The professor advised her student on his paper yesterday.*
How long did this take?

a) up to 25 minutes
b) between 25 minutes and 1 hour
c) between 1 hour and 2 weeks
d) more than 2 weeks

The same lists as in the previous studies were used.

3.2 Results

Figure 4 shows the proportional distribution of categorizations into the four time bins, and the proportion of counts in each category, depending on event category and grammatical construction. In general, we observed a tendency to categorize all events into the shorter time bins. More theoretically crucial, light verb constructions were categorized as being shortest.

Data were analyzed using R with a mixed-effects cumulative logit model using the `ordinal` package (Christensen, 2015). For ordered response categories (in our case, four bins ranging from shortest to longest), a mixed-effects cumulative logit model specifies response probabilities for a given data point as a function of predictor variables. Instead of the intercept, ordered logit models provide a set of threshold parameters, which describe the boundaries from one bin to the next, and the probability of being drawn from one particular bin is estimated by the linear predictors with the inverse logit function (see Appendix B for an in-depth explanation of the mathematical underpinnings). The predictor variables in our model were construction (transitive or ditransitive), event category (PC, DC, or DM), and their interaction. We Helmert-coded predictors and used maximal random effects structure as in Experiment 1a. Under this coding, construction is coded with transitive=–1, ditransitive=1. Response categories were coded 1 to 4 in order of increasing duration.

Table 4 shows the statistical results of the categorization task. The first three rows indicate the threshold coefficients from one bin into the next. The middle part of the table shows the regression coefficient for construction – the –0.195 value indicating that the ditransitive construction is associated with shorter event durations – and the results of the likelihood-ratio tests for the main effects and their interaction: Analogous to the analyses in Experiment 1a and 1b, we separately tested for main effects of construction and event category, comparing models with or without a fixed effect of event category or construction, respectively, to the same model with both fixed effects and their interaction. Construction had a significant effect on duration estimates in the model comparison, and the interaction between construction and event category was marginally significant, but the main effect of event category was not. The lack of a main effect of event category is reassuring given that the time bins were constructed based on duration estimates from Experiment 1a.

Finally, the bottom part of Table 4 displays the results of likelihood-ratio tests for each event category, testing the main effect of construction only. 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, a look at the β-estimates tells us that the effect of construction went in opposite directions for durative count items, compared to punctive count and durative mass items: 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.

3.3 Discussion of Experiment 2

This study gave participants the opportunity to estimate event durations using predefined time bins as choices, which might have been easier for them than coming up with time estimates on their own. The answer options in Experiment 2 were based on each item pair's averaged estimates from Exper-

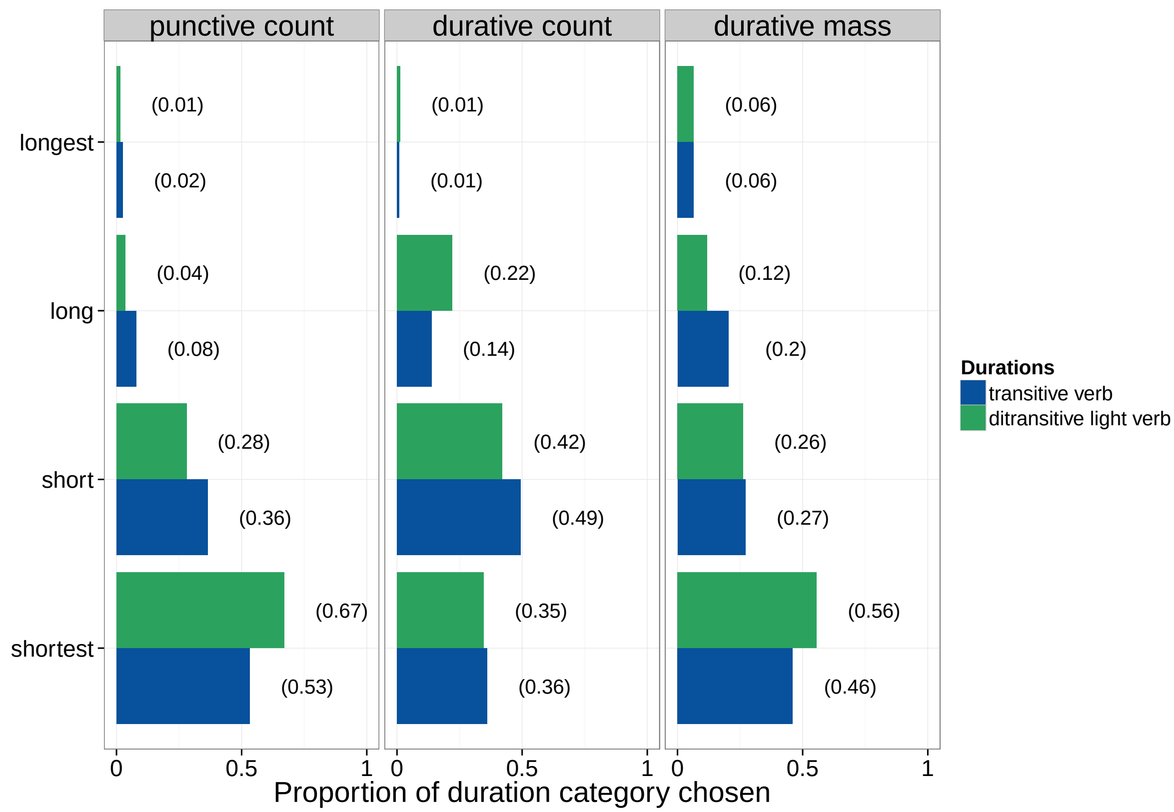

Figure 4: For PC and DM items, participants chose the shortest durations in ditransitive light verb frame, and longer durations in the transitive verb frames. There was no significant difference for DC items.

	β	SE		
shortest|short	-0.114	0.256		
short|long	2.201	0.265		
long|longest	4.669	0.320		

	Df	β	LR.stat	p-value
construction	1	-0.195	4.079	0.043 *
event category	2		4.46	0.107 *n.s.*
construction × event category	2		5.21	0.070 .
PC – construction	1	0.811	4.995	0.025 *
DC – construction	1	-0.289	1.263	0.262 *n.s.*
DM – construction	1	0.607	2.156	0.142 *n.s.*

Table 4: Experiment 2: Regression table for categorizations.

iment 1a, which, as we discussed above, seemed fairly high. Nevertheless, we found significant differences for PC events: When presented in ditransitive light verb constructions *(give a kiss)*, they were estimated to last less time then in transitive syntax *(kiss)*. For DC and DM events *(talking/giving a talk, advising/giving advice)*, we did not observe such a difference, although the β-estimates indicate that DM events follow the same trend as PC events, with ditransitive structures pushing categorizations towards shorter bins.

Interestingly, the largest proportion of choices in both grammatical constructions and all three event types fell to the shortest options. Given that the choice options were based on previously obtained quartiles, the answers should have been roughly equally distributed. Thus, our data contribute to the literature on over- and underestimation of event duration in an interesting way: Open guesses as in Experiment 1a and 1b tend to overestimate (at least for the event types we used here, which, unlike in classical studies on the planning fallacy, did not include unpleasant chore-like events), whereas predefined categorizations are closer to more realistic event durations.[5]

As stated above, though, we were not *per se* interested in how long events are estimated to take, but in the influence of mass and count syntax on estimated event durations. Crucially, as in Experiment 1a and 1b, we found

[5]We tested this explanation by giving participants shorter answer options that were not based on previous estimates. We found a much more equal distribution of answers. Since this study is tangential to the goal of this paper, it is not reported in detail, but the data are accessible under [**repository address anonymous for peer review**].

that count syntax shortens the time estimate for punctive, but not durative events. We had predicted this pattern because punctive verbs are often interpreted iteratively, and by using count syntax, one picks out one particular subevent. For durative verbs with count syntax, however, there is no distinct subevent to be picked out: Thus, we hypothesized that durative verbs might undergo a larger conceptual shift when when put in count syntax *(to give a talk)* than do punctive count or durative mass verbs.

The next three studies test these explanations: Experiments 3a and 3b investigate how many event repetitions are imagined by comprehenders, depending on the syntactic construction; and Experiment 4 asks whether using mass and count syntax introduced a conceptual shift between transitive and ditransitive mass versus count syntax.

4 Experiment 3a: Event repetition

This section presents two studies that establish whether using count or mass syntax affects how many times an event is understood as occurring.

The theoretically most crucial prediction applies to punctive events. Punctive events, like kissing, are often understood iteratively, even if their lexical semantics merely conveys one single, bounded, event (Kim and Kaiser, 2015). For example, people might interpret *Douglas kissed Mary* as him kissing her more than once. However, when presented in count syntax (*to give a kiss*), only one kiss should be singled out. This would explain the consistently lower time estimates obtained in Experiments 1a, 1b, and 2.

Obviously, in order to be counted, a given event needs to be individuated, and we expect individuation to be easier in count than in mass syntax (Barner, Wagner, and Snedeker, 2008). So, for durative events in mass syntax, predictions are not quite as straightforward: *The professor advised her student* consists of many individual advising situations. The bounded verb *give* in *The professor gave advice to her student* introduces a boundary to the advising event (Krifka, 1992); however, no definite article aids in the event individuation, so we might expect a weaker effect than for the punctive count events.

However, for durative events that can enter count syntax, we might predict the same trend as for punctive count events: *talking* could convey more events than *giving a talk*. Note that this prediction is independent of our second prediction, namely that there is a conceptual shift involved from a transitive durative verb to the same lexical item in ditransitive count syntax; and that any changes in event repetition counts would still be orthogonal to changes in event duration.

4.1 Methods

4.1.1 Participants

For this study, we recruited 80 self-described English native speakers from Amazon Mechanical Turk.

4.1.2 Procedure

Participants read each sentence and then noted how many events they imagined reading the sentence. The exact instructions were:

> "In this survey, we would like to know how many events you imagine when reading sentences. For example, when I say, "Tim flew from Boston to Baltimore", you probably imagine one flight (perhaps two, if you're pretty sure that he had to change planes). But when I say, "Lisa sniffled", you might not imagine one sniffle, but maybe three, or four, or five, because that's how sniffling usually goes."

4.1.3 Stimuli

We used the stimuli described in Table 1, counterbalanced across two lists.

4.2 Results

Our data spanned several orders of magnitude, (e.g., responses for punctive events in transitive frame ranged from 1 to 60), effects of construction are likely to operate proportional to intrinsic construal of numbers of events, as was the case with construed event durations in Experiments 1a and 1b. Therefore we log-transformed event-count responses for purposes of data summarization and analysis.

Mean count of events was lower for ditransitive light verb constructions in each event category. Figure 5 shows the pattern of results, with log event

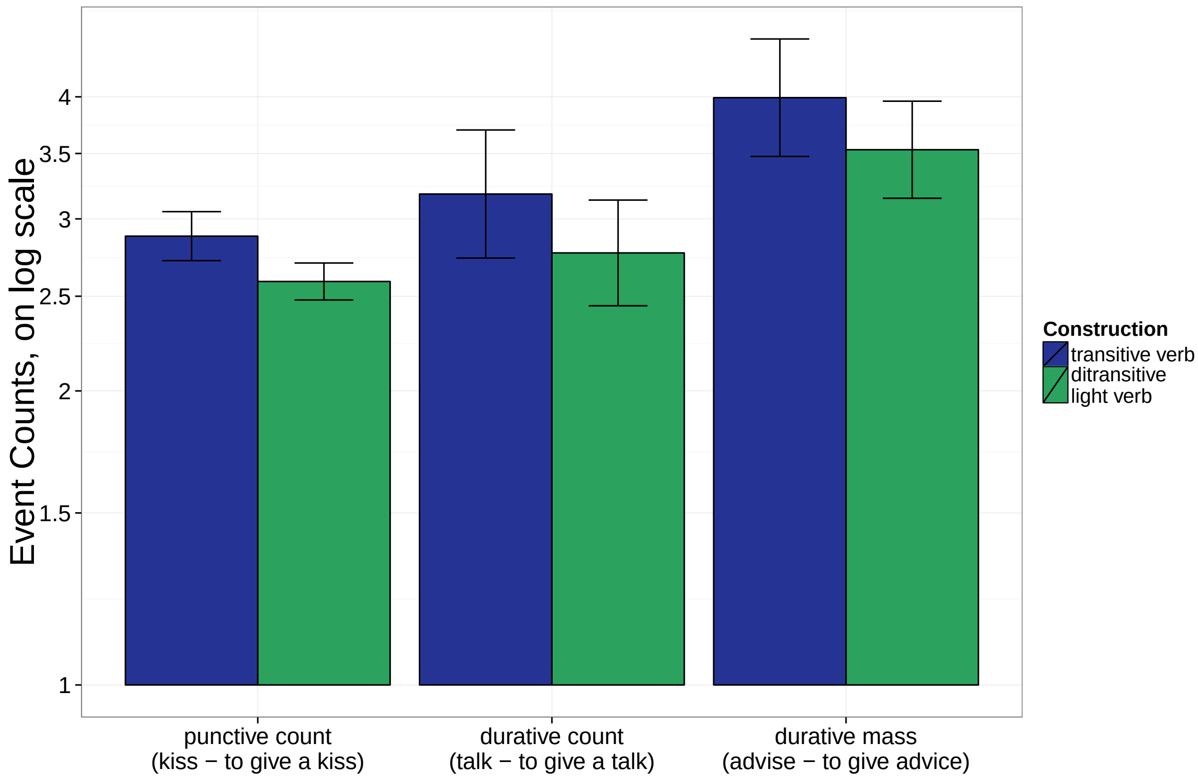

Figure 5: Count of imagined events in Experiment 3a. The y-axis is represented in log scale, but labeled with non-transformed counts for convenience. SEs represent 95% confidence interval.

counts back-transformed to raw event counts for convenience of interpretation: 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. These effects were also numerically present for durative count and durative mass items, but variances in these constructions were higher.

As in Experiment 1, we used linear mixed-effects regression models with maximal random effects structure for statistical analysis. Again, we included event category (PC, DC, and DM), construction (transitive or ditransitive light verb) and their interaction as fixed effects, random by-participant intercepts and slopes for all fixed effects, and random by-item intercepts and slopes for event category, using the `lme4` package. We used Helmert coding for both fixed-effects predictors. For event category, we grouped PC and DM together as one Helmert contrast pair, and their average contrasted with DC items as the second Helmert contrast pair.

	Df	χ^2	p-value
construction	1	7.011	0.008 **
event category	2	5.326	0.069 .
construction × event category	2	0.016	0.992 *n.s.*
PC – construction	1	7.187	0.007 **
DC – construction	1	1.086	0.297 *n.s.*
DM – construction	1	1.030	0.310 *n.s.*

Table 5: Table of likelihood estimation results for event counts in Experiment 3, testing the main effects and their interaction (upper part), and pairwise comparisons (lower part).

As in the previous experiments, we separately tested for main effects of construction and event category with model comparisons. The comparisons revealed that construction was a significant main effect on event counts, event category was marginally significant, and their interaction was not. This is reflected in Table 5: The upper part displays the full mixed regression model with construction and event category and their interaction. With punctive transitive events as intercept, there was a main effect of construction, and an interaction between construction and durative mass events, which is also reflected in the pairwise models (lower part).

4.3 Replication: Experiment 3b

We replicated this study by asking 80 native speakers on Amazon Mechanical Turk for the specific event types after each sentences, for example "How many kisses did you just imagine?", "How many talks did you just imagine?", or "How many advices did you just imagine?". Note that predictions for durative count items might be stronger in this study than in Experiment 3a, since we asked for the specific event. We obtained a similar pattern of results (see Figure 6).

Table 6 displays the results of the event count estimation test for the main effects, which were both significant, and their interaction, which was not. Pairwise comparisons (bottom part of Table 6) show main effects of

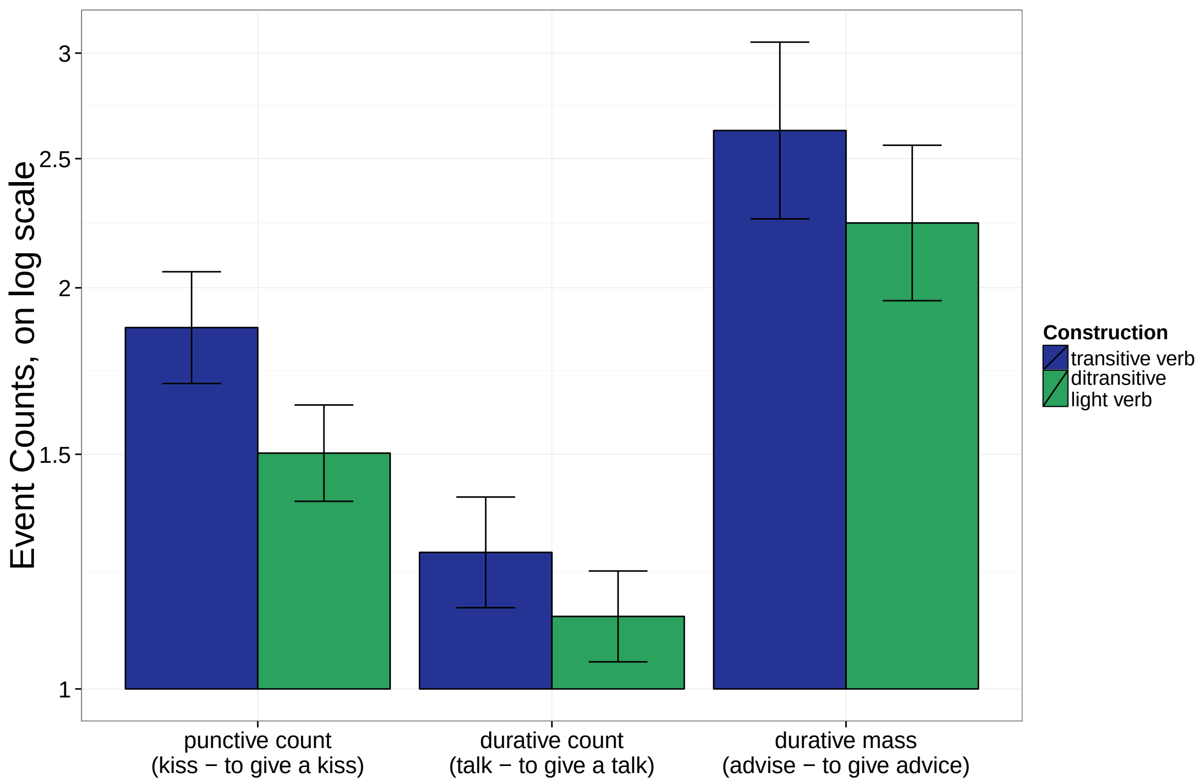

Figure 6: Count of imagined events in Experiment 3b. The y-axis is represented in log scale, but labeled with non-transformed counts for convenience. SEs represent 95% confidence interval.

construction for both punctive count and durative count, but not for durative mass events.

	Df	χ^2	p-value
construction	1	11.034	0.000 ***
event category	2	8.297	0.016 *
construction × event category	2	1.001	0.606 *n.s.*
PC – construction	1	5.391	0.020 *
DC – construction	1	4.063	0.043 *
DM – construction	1	2.234	0.135 *n.s.*

Table 6: Table of likelihood estimation results for event counts in Experiment 3b, testing the main effects and their interaction (upper part), and pairwise comparisons (lower part).

4.4 Discussion of Experiment 3a and 3b

This study investigated whether phrasing an event with mass or count syntax, instead of with a transitive verb, affects any iterative readings that are present.

Crucially, we see a significant reduction in imagined number of events as a measure of implicit iterativity from transitive verb encoding to ditransitive light verb encoding – but, as predicted, only consistently in punctive count events, and to a lesser degree, in durative count events. This, the more pronounced effect for durative count items in Experiment 3b, and the overall reduction in event counts in Experiment 3b, confirm our hypothesis that count syntax encourages individuating over subevents, leading to shorter event conceptualization.

Contrary to what the previous literature has claimed, durative events resulted in iterative readings, as well; we attribute that to the fact that events like *advising* or *talking*, while not easily segmentable by pieces of advice or specific identifiable talking events, do carry an element of interactivity: There is a back-and-forth between the advisor or talker, and the addressee; and there might also be habitual interpretations available. Thus, it could be that our participants conceptualized the "number of events" question as "number of subevents" – which would explain the occasional "more than one" answer for count syntax like *give a talk*.

However, we had hypothesized that count syntax might encourage a conceptual shift in durative verbs that acts orthogonally to any changes in temporal conceptualization. The last experiment investigates this possibility.

5 Experiment 4: Event Similarity

This study tested another prediction of our theory: Durative events in count syntax (*to give a talk*) are conceptually further apart from their transitive verb counterparts (*to talk*) than punctive events in count syntax, or durative events in mass syntax. This prediction is drawn from the analogy to mass nouns, such as *glass* or *iron* into count syntax: In many cases, the count noun denotes objects that are conceptually related, yet different, from the noun in mass syntax, such as in *a glass* or *an iron*. If the durative events in count syntax behave in parallel, we should expect them to be conceptually further apart from their transitive counterparts than punctive events in count or durative events in mass syntax.

5.1 Methods

5.1.1 Participants

We recruited 40 self-described English native speakers from Amazon Mechanical Turk for this study.

5.1.2 Procedure

We asked participants to rate event similarity between transitive and ditransitive frames on a 7-point Likert scale where 1 indicated "same event", and 7, "completely different event". The exact instructions were:

> "In this survey, we would like to know how similar are events to each other on a scale from "exactly the same" to "completely different". For example, "sleeping on the couch" and "sleeping on the sofa" are exactly the same event, so in the questionnaire below, you would click on the very first button. "Saluting the king" and "bowing to the king", on the other hand, are somewhat similar events, so you would probably judge that to be in the middle. Finally, "eating ice cream" and "driving a car" are

completely different events, so you would click on the rightmost button."

5.1.3 Stimuli

We used the 20 item pairs shown in Table 1, without the sentence context. In addition, we created 26 filler pairs that ranged from very close synonyms (*repairing – fixing*) to very different events (*working – being lazy*).

5.2 Results

For filler items, the average rating was 3.62 (SD=2.1), with the full range of the scale being used.

Figure 7 shows the rating results for critical pairs. Punctive events (*to kiss* vs *to give a kiss*) received a mean rating of 1.6 (SD=1.0). Pairs of durative verbs and durative verbs in mass syntax (*to advise* vs *to give advice*) received a similarly low rating (mean: 1.5, SD=.9). This means that for both punctive count and durative mass items, both constructions denote very similar events. As predicted, difference ratings for pairs of durative verbs and durative verbs in count syntax (*to talk* vs *to give a talk*) were higher: 2.1 (SD=1.5).

	Df	χ^2	p-value
event category	2	6.819	0.033 *
PC vs DC	1	3.799	0.046 *
DC vs DM	1	4.869	0.027 *
PC vs DM	1	0.000	0.984 *n.s.*

Table 7: Table of likelihood estimation results for event similarity in Experiment 4, testing the main effect of event category (upper part), and pairwise comparisons between event categories (lower part).

Like in Experiment 1a and 1b, we used comparisons among linear mixed models to test for a main effect. Both the model without and the model with event category as fixed-effect predictor included random intercepts for items and participants, as well as random by-participant slopes for event category. As the model comparisons show, event category was a significant main effect for similarity ratings. Table 7 displays the results from the full model, as well as pairwise models comparing all event categories with each other. As

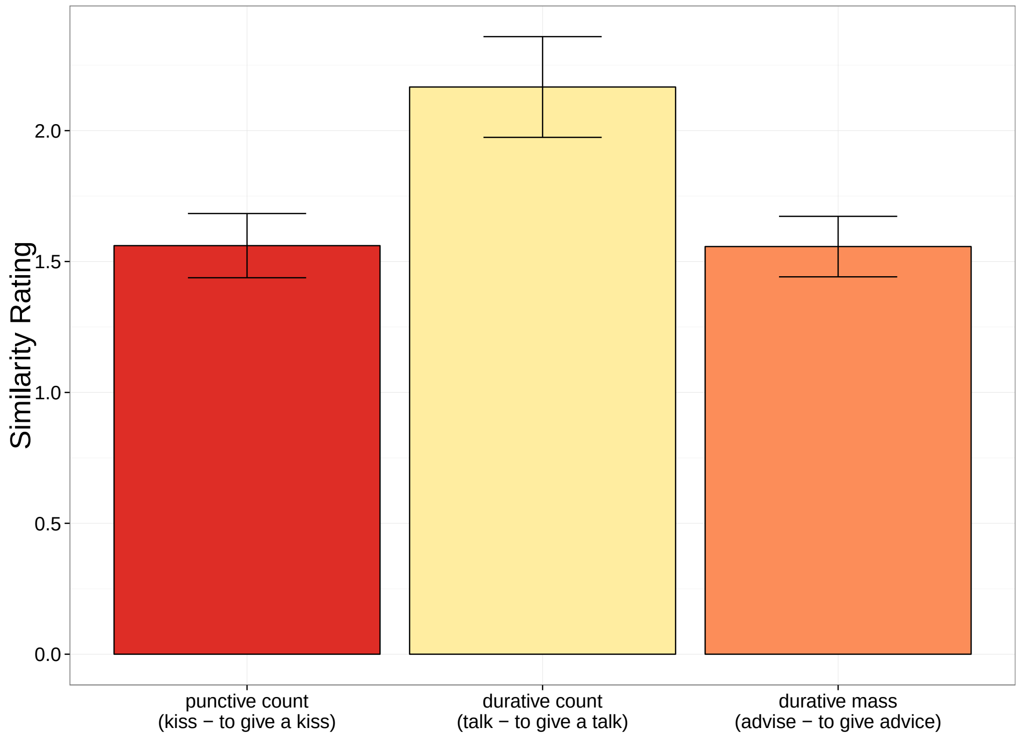

Figure 7: PC and DM item pairs were rated more similar to each other than DC item pairs.

the data show, there was no significant difference between similarity ratings for punctive count and durative mass event pairs; durative count pairs, on the other hand, were rated significantly less similar to each other than pairs in the other two event categories.

5.3 Discussion of Experiment 4

Experiment 4 confirms the intuition that while *giving a kiss* and *kissing*, as well as *giving advice* and *advising*, belong ontologically to the same event, *giving a talk* and *talking* are conceptually further apart – albeit with considerable overlap. This result contributes to claims made in the literature that light verb constructions often help to close lexical gaps (Glatz, 2006; Grimshaw and Mester, 1988; Miyagawa, 1989), and it draws our attention to interesting parallels to using count syntax on mass nouns (Gordon, 1985; Srinivasan and Rabagliati, 2015; Wiese and Maling, 2005).

These results are crucial to our question of how using mass or count syntax affects the construal of event duration: If a durative lexeme in count syntax covers a conceptually different event than the same lexical item in a transitive verb frame, the conceptualization of event duration applies to the new concept, and thus any changes in duration estimates would be largely coincidental.

6 General Discussion

We have presented a family of experiments that tested the hypothesis that describing an event with mass versus count syntax affects the construal of event similarity and duration in a way that is systematically predictable from the interaction of mass versus count syntax and verb semantics (see Figure 1 on page 4). These predictions were built on insights from the semantics literature that pointed out similarities between mass nouns and durative events on the one side, and count nouns and punctive events on the other side (Harley, 2005; Jackendoff, 1991; Krifka, 1992).

Specifically, we had predicted that punctive events in count syntax (*give a hug*) are construed as taking less time than in transitive verb frame (*hug*); this pattern was predicted because in punctive events, which are often interpreted iteratively (Kim and Kaiser, 2015), one subevent is singled out by count syntax.

In durative events, the same pattern was predicted for events in mass syntax, such as *giving advice*. In combination with the bounded verb *give*, the mass noun carves out a limited portion of the event structure, leading to shorter event conceptualization.

A different pattern was predicted for durative verbs in count syntax (*give a talk* versus *talk*): We expected durative events in count syntax to be semantically further from their verbal counterparts (*to talk – a talk*) than punctive events (*to kiss – a kiss*); since these shifts were presumable orthogonal to the temporal structure of the event, we did not make any predictions about duration estimates. These predictions stemmed from insights about the parallels between durative events and mass nouns: When some mass nouns are forced into count syntax (*a glass; an iron*), there is a semantic shift from the mass noun meaning (*glass; iron*).

Our experimental results broadly confirm these predictions. In Experiments 1a and 1b, we elicited open duration estimates, which were consistently lower for punctive events in count syntax and durative events in mass syntax than when they occurred in transitive frames; durative events in count syntax did not show this effect. Experiment 2 replicated these results by using the quartiles of each event's individual duration estimates obtained in Experiment 1a as answer choices; again, punctive count and durative mass ditransitive structures were judged to take less time than transitive structures, while there was no difference for durative events in count syntax.

On our theory, a key factor for the reduction of duration estimates in punctive verbs was that people should imagine fewer events taking place than in the ditransitive count frame, due to the combination of a bounded verb (*give*) and nominal count syntax (*a kiss*). However, this effect should be weaker for durative events in either count or mass syntax when participants are asked for number of events. Experiment 3a confirmed these predictions. Experiment 3b showed that when people are asked for how many specific events they imagined – how many talks, for example – the reduction in event counts is significant even for durative events in count syntax. This confirms both the observation that event individuation is easier in count syntax (Barner, Wagner, and Snedeker, 2008), and that durative verbs behave similarly to mass nouns (Krifka, 1992).

Experiment 4 showed, consistent with our predictions, that durative events undergo a semantic shift in count syntax: The semantic differences between transitive and ditransitive frames were larger in durative count pairs (*talk – give a talk*) than in punctive-count pairs (*hug –give a hug*) or durative-mass pairs (*advise – give advice*).

Thus, the shift from a transitive to a ditransitive frame has systematically predictable repercussions, depending on whether the event type was durative or punctive, and depending on whether the event was described with a mass or with a count noun.

These results provide psycholinguistic evidence for the observation in formal semantics that reference properties of syntactic objects change the reference properties of the whole predicate (Krifka, 1992; also see Quine, 1969; Verkuyl, 1972). In our case, the nominalization of eventive verbs using light verb constructions with a bounded verb helped to divide experience into countable units (for punctive verbs in count syntax and durative verbs in mass syntax), or introduced a semantic shift (for durative verbs in count syntax), similar to what happens with mass nouns in count syntax.

Our results also complement studies that suggest that people conceptualize events differently depending on subtle choices among syntactic alternations (Fausey and Boroditsky, 2010; Johnson and Goldberg, 2013; Wittenberg and Snedeker, 2014) and that conversely, subtle changes in event structure result in changing preferences between syntactic alternations (Gropen et al., 1991). This growing family of experiments, together with a number of corpus studies (Benor and Levy, 2006; Bresnan et al., 2007), contributes to our understanding how syntactic and semantic structures, as well as processing pressures, lead to a speaker's decision between to seemingly equivalent constructions.

These findings may also help interpret results from less explicit tasks. We know from behavioral, ERP, and MEG studies that light verb constructions are processed differently from non-light constructions (Briem et al., 2009; Piñango, Mack, and Jackendoff, in press; Wittenberg and Piñango, 2011; Wittenberg et al., 2014). So far, however, only semantic role mismatches had been identified as a factor contributing to the processing difference (Wittenberg, Khan, and Snedeker, under review; Wittenberg and Snedeker, 2014). Based on the present work, we might hypothesize that the calculation of temporal structure also plays a role in the real-time processing of light verb constructions.

A question this paper raises is how well these results might generalize to other constructions such as idioms, or to other light verb constructions with unbounded verbs. In the literature on idiom comprehension there is a consensus that the lexical items in an idiomatic phrase such as *to kick the bucket* are accessed individually as well as holistically (Cutting and Bock, 1997; Holsinger, 2013; Sprenger, Levelt, and Kempen, 2006; Tabossi, Wolf, and Koterle, 2009; Titone and Libben, 2014). So if an idiom consists of a semelfactive verb (*kick*) and a count noun (*bucket*), and its semantically transparent counterpart is a semantic achievement (*to die*) we would predict that *John kicked the bucket* will be estimated to take a shorter time than *John died*. Similarly, we would predict that other light verb constructions with bounded verbs would lead to the same effect; for example, *John took a shower* should be estimated as taking less time than *John showered*, whereas *John had a shower* should affect estimated event durations to a lesser degree, since both verbs (*to shower* and *to have*) are unbounded. These predictions remain to be addressed in future studies.

In sum, this paper showed that using mass versus count syntax affects the construal of event similarity and duration in a way that is systematically predictable from the interaction of mass versus count syntax and verb semantics. Our results confirm hypotheses from formal semantics and psycholinguistics, and connect to the literature on duration estimates in neighboring fields, such as behavioral economics. Thus, these studies advance our understanding of the effects of syntactic choices on subtle aspects of event construal.

References

Allerton, D. (2002). *Stretched Verb Constructions in English*. London.

Baayen, R Harald, Douglas J Davidson, and Douglas M Bates (2008). "Mixed-effects modeling with crossed random effects for subjects and items". In: *Journal of Memory and Language* 59.4, pp. 390–412.

Barner, David and Alan Bale (2002). "No nouns, no verbs: psycholinguistic arguments in favor of lexical underspecification". In: *Lingua* 112.10, pp. 771–791.

Barner, David and Jesse Snedeker (2006). "Children's early understanding of mass-count syntax: individuation, lexical content, and the number asymmetry hypothesis". In: *Language Learning and Development* 2.3, pp. 163–194.

Barner, David, Laura Wagner, and Jesse Snedeker (2008). "Events and the ontology of individuals: Verbs as a source of individuating mass and count nouns". In: *Cognition* 106.2, pp. 805–832.

Barr, Dale J. et al. (2013). "Random effects structure for confirmatory hypothesis testing: Keep it maximal". In: *Journal of Memory and Language* 68.3, 255–278.

Bates, Douglas et al. (2014). "Package 'lme4'". In: *R Foundation for Statistical Computing, Vienna*.

Benor, Sarah Bunin and Roger Levy (2006). "The chicken or the egg? A probabilistic analysis of English binomials". In: *Language*, pp. 233–278.

Boltz, Marilyn G (1995). "Effects of event structure on retrospective duration judgments". In: *Perception & psychophysics* 57.7, pp. 1080–1096.

Bresnan, Joan et al. (2007). "Predicting the dative alternation". In: *Cognitive foundations of interpretation*, pp. 69–94.

Briem, Daniela et al. (2009). "Distinct processing of function verb categories in the human brain". In: *Brain Research* 1249, pp. 173–180.

Brugman, Claudia (2001). "Light verbs and polysemy". In: *Language Sciences* 23.4, pp. 551–578.

Buhrmester, Michael, Tracy Kwang, and Samuel D Gosling (2011). "Amazon's Mechanical Turk a new source of inexpensive, yet high-quality, data?" In: *Perspectives on psychological science* 6.1, pp. 3–5.

Burt, Christopher DB and Jennifer S Popple (1996). "Effects of implied action speed on estimation of event duration". In: *Applied cognitive psychology* 10.1, pp. 53–63.

Butt, Miriam (2003). "The Light Verb Jungle". In: *Harvard Working Papers in Linguistics* 9, pp. 1–49.

— (2010). "The light verb jungle: still hacking away". In: *Complex Predicates in Cross-Linguistic Perspective*. Ed. by B. Baker M. Amberber M. Harvey. Cambridge University Press, pp. 48–78.

Christensen, R. H. B. (2015). *ordinal—Regression Models for Ordinal Data*. R package version 2015.1-21. http://www.cran.r-project.org/package=ordinal/.

Crump, Matthew JC, John V McDonnell, and Todd M Gureckis (2013). "Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research". In: *PloS one* 8.3, e57410.

Cutting, J Cooper and Kathryn Bock (1997). "That's the way the cookie bounces: Syntactic and semantic components of experimentally elicited idiom blendsß". In: *Memory & Cognition* 25.1, pp. 57–71.

Dowty, David (1991). "Thematic Proto-Roles and Argument Selection (Prepublication version)". In: *Language* 67:3, pp. 547–619.

Fausey, Caitlin M and Lera Boroditsky (2010). "Subtle linguistic cues influence perceived blame and financial liability". In: *Psychonomic Bulletin & Review* 17.5, pp. 644–650.

Flecken, Monique and Johannes Gerwien (2013). "Grammatical aspect modulates event duration estimations: findings from Dutch". In: *the 35th Annual Meeting of the Cognitive Science Society (CogSci 2013)*. Cognitive Science Society, pp. 2309–2314.

Garrod, Simon and Martin J. Pickering (Dec. 2003). "Linguistics fit for dialogue". In: *Behavioral and Brain Sciences* 26 (06), pp. 678–678.

Glatz, Daniel (2006). "Funktionsverbgefüge–semantische Doubletten von einfachen Verben oder mehr". In: *Domänen der Lexikalisierung kommunikativer Konzepte* 33, p. 129.

Goldberg, Adele (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. University of Chicago Press.

— (2003). "Constructions: a new theoretical approach to language". In: *Trends in Cognitive Sciences* 7.5, pp. 219–224.

Gordon, Peter (1985). "Evaluating the semantic categories hypothesis: The case of the count/mass distinction". In: *Cognition* 20.3, pp. 209–242.

Grimshaw, Jane and Arnim Mester (1988). "Light Verbs and Theta-Marking". In: *Linguistic Inquiry* 19, pp. 205–232.

Gropen, Jess et al. (1991). “Affectedness and direct objects: The role of lexical semantics in the acquisition of verb argument structure”. In: *Cognition* 41.1, pp. 153–195.

Hale, Kenneth and Jay Keyser (1993). “On argument structure and the lexical expression of syntactic relations”. In: *The View from Building 20*. Ed. by Kenneth Hale and Jay Keyser. Cambridge: MIT Press, pp. 53–109.

Halkjelsvik, Torleif, Magne Jørgensen, and Karl Halvor Teigen (2011). “To read two pages, I need 5 minutes, but give me 5 minutes and I will read four: How to change productivity estimates by inverting the question”. In: *Applied Cognitive Psychology* 25.2, pp. 314–323.

Halle, Morris and Alec Marantz (1994). “Some key features of Distributed Morphology”. In: *MIT working papers in linguistics* 21.275, p. 88.

Harley, Heidi (2005). “How do verbs get their names? Denominal verbs, manner incorporation, and the ontology of verb roots in english”. In: *The Syntax of Aspect*. University Press, pp. 42–64.

Holsinger, Edward (2013). “Representing idioms: Syntactic and contextual effects on idiom processing”. In: *Language and speech* 56.3, pp. 373–394.

Jackendoff, Ray (1974). “A Deep Structure Projection Rule”. In: *Linguistic Inquiry* 5.4, pp. 481–505.

Jackendoff, Ray (1991). “Parts and boundaries”. In: *Cognition* 41.1, pp. 9–45.

— (2002). *Foundations of Language*. Oxford University Press.

— (2007). “A Parallel Architecture perspective on language processing”. In: *Brain Research* 1146, pp. 2–22.

Jespersen, Otto (1954). *A Modern English Grammar on Historical Principles*. London: Allen & Unwin and Copenhagen: Ejnar Munksgaard.

Johnson, Matt A. and Adele E. Goldberg (2013). “Evidence for automatic accessing of constructional meaning: Jabberwocky sentences prime associated verbs”. In: *Language and Cognitive Processes* 28.10, pp. 1439–1452.

Kahneman, Daniel and Amos Tversky (1977). *Intuitive prediction: Biases and corrective procedures*. Tech. rep. DTIC Document.

Kim, L.K. and E. Kaiser (2015). “The number of times a motion repeats influences sentence processing”. In: *Proceedings of the Cognitive Science Society Conference*.

Krifka, Manfred (1992). “Thematic relations as links between nominal reference and temporal constitution”. In: *Lexical Matters*. Ed. by Ivan Sag and Anna Szabolcsi. CSLI Publications, Chicago University Press, pp. 29–53.

Kruger, Justin and Matt Evans (2004). “If you don’t want to be late, enumerate: Unpacking reduces the planning fallacy”. In: *Journal of Experimental Social Psychology* 40.5, pp. 586–598.

Levy, Roger (2008). “Expectation-Based Syntactic Comprehension”. In: *Cognition* 106.3, 1126–1177.

— (2013). “Memory and Surprisal in Human Sentence Comprehension”. In: *Sentence Processing*. Ed. by Roger P. G. van Gompel. Hove: Psychology Press, 78–114.

— (2014). “Using R formulae to test for main effects in the presence of higher-order interactions”. In: *arXiv:stat.ME/1405.2094v1*.

Loftus, Elizabeth F et al. (1987). “Time went by so slowly: Overestimation of event duration by males and females”. In: *Applied Cognitive Psychology* 1.1, pp. 3–13.

Macar, Françoise, Simon Grondin, and Laurence Casini (1994). “Controlled attention sharing influences time estimation”. In: *Memory & Cognition* 22.6, pp. 673–686.

Madden, Carol J and Rolf A Zwaan (2003). “How does verb aspect constrain event representations?” In: *Memory & Cognition* 31.5, pp. 663–672.

Magliano, Joseph P and Michelle C Schleich (2000). “Verb aspect and situation models”. In: *Discourse processes* 29.2, pp. 83–112.

Miyagawa, Shigeru (1989). “Light verbs and the ergative hypothesis”. In: *Linguistic inquiry*, pp. 659–668.

Newman, John (1996). *Give: A cognitive linguistic study*. Vol. 7. Walter de Gruyter.

Paczynski, M, R Jackendoff, and G Kuperberg (2014). “When Events Change Their Nature: The Neurocognitive Mechanisms Underlying Aspectual Coercion.” In: *Journal of Cognitive Neuroscience*.

Pedersen, Alice CI and Daniel B Wright (2002). “Do differences in event descriptions cause differences in duration estimates?” In: *Applied Cognitive Psychology* 16.7, pp. 769–783.

Piñango, Maria M., Jennifer Mack, and Ray Jackendoff (in press). “Semantic Combinatorial Processes in Argument Structure: Evidence from Light-Verbs”. In: *Proceedings of Berkeley Linguistics Society*.

Quine, Willard V (1969). “Word and object”. In: *Cambridge, Mass.*

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria.

Roy, Michael M, Nicholas JS Christenfeld, and Craig RM McKenzie (2005). "Underestimating the duration of future events: memory incorrectly used or memory bias?" In: *Psychological bulletin* 131.5, p. 738.

Sprenger, Simone A., Willem J.M. Levelt, and Gerard Kempen (Feb. 2006). "Lexical access during the production of idiomatic phrases". In: *Journal of Memory and Language* 54.2, pp. 161–184.

Srinivasan, Mahesh and Hugh Rabagliati (2015). "How concepts and conventions structure the lexicon: Cross-linguistic evidence from polysemy". In: *Lingua*.

Tabossi, Patrizia, Kinou Wolf, and Sara Koterle (2009). "Idiom syntax: Idiosyncratic or principled?" In: *Journal of Memory and Language* 61.1, pp. 77–96.

Titone, Debra and Maya Libben (2014). "Time-dependent effects of decomposability, familiarity and literal plausibility on idiom priming: A cross-modal priming investigation". In: *The Mental Lexicon* 9.3, pp. 473–496.

Tse, Peter Ulric et al. (2004). "Attention and the subjective expansion of time". In: *Perception & psychophysics* 66.7, pp. 1171–1189.

Vendler, Zeno (1957). "Verbs and times". In: *The philosophical review*, pp. 143–160.

Verkuyl, Hendrik Jacob (1972). *On the compositional nature of the aspects*. Vol. 15. Springer.

Wiese, Heike (2006). ""Ich mach dich Messer" – Grammatische Produktivität in Kiez-Sprache ("Kanak Sprak")". In: *Linguistische Berichte* 207, pp. 245–273.

Wiese, Heike and Joan Maling (2005). "Beers, kaffi, and Schnaps – Different grammatical options for 'restaurant talk' coercions in three Germanic languages". In: *Journal of Germanic Linguistics* 17:1, pp. 1–38.

Wittenberg, Eva, Manizeh Khan, and Jesse Snedeker (under review). "Investigating thematic roles through implicit learning."

Wittenberg, Eva and Maria Mercedes Piñango (2011). "Processing light verb constructions". In: *The Mental Lexicon* 6.3, pp. 393–413.

Wittenberg, Eva and Jesse Snedeker (2014). "It takes two to kiss, but does it take three to give a kiss? Categorization based on thematic roles". In: *Language, Cognition and Neuroscience* 29.5, pp. 635–641.

Wittenberg, Eva et al. (2014). "The difference between "giving a rose" and "giving a kiss": Sustained neural activity to the light verb construction". In: *Journal of Memory and Language* 73, pp. 31–42.

Appendix A: Stimuli used for Experiments 1-3

Stimuli are shown in transitive frame. In the ditransitive light verb construction, the underlined verb would serve as the light noun, i.e., *After their first date, Douglas kissed Mary* $\rightarrow$ *After their first date, Douglas gave a kiss to Mary.*

1. Punctive Count
   - (a) *After their first date, Douglas kissed Mary.*
   - (b) *Laughing nastily, the thug kicked the victim.*
   - (c) *When they met up, Nathasha hugged Cynthia.*
   - (d) *Noam embraced Jennifer before they split up.*
   - (e) *Julius cuddled his brother at bedtime.*
   - (f) *Laura poked Owen because he was so annoying.*
   - (g) *Martin shook the cocktail before he served it.*
2. Durative Count
   - (a) *The professor lectured on injustice yesterday.*
   - (b) *The CEO talked about his latest sales strategy last night.*
   - (c) *The President spoke about affordable education on Wednesday.*
   - (d) *During the community art show, the composer presented his latest work.*
   - (e) *After the graduation ceremony, Tom Hanks addressed the students.*

(f) *The mother scolded the child today.*

3. Durative Mass

   (a) *The professor advised her student on his paper yesterday.*

   (b) *Yesterday, the nurse assisted Dr. Kohler in the emergency room.*

   (c) *After the job interview, Sheila assured Keith.*

   (d) *Before Kelly left for a year abroad, her friends encouraged her.*

   (e) *After the devastating flood, the mayor recognized the rescue workers.*

   (f) *His new girlfriend supported Sam after his divorce.*

   (g) *At the summer party, the university official thanked the academic community for their efforts.*

Appendix B: The mixed-effects cumulative logit model used for Experiment 2

Data were analyzed using R (R Core Team, 2014), specifically with mixed-effects cumulative logit model using the `ordinal` package (Christensen, 2015). For ordered response categories $1, 2, \ldots, N$, a mixed-effects cumulative logit model specifies response probabilities for a given datum i as a function of predictor variables as follows. There are $N-1$ linear predictors, each of the following form:

$$\eta_{ij} = \alpha_j + \sum_k \beta_k x_{ik} + \sum_k b_k z_{ik} \qquad j \in \{1, 2, \ldots, N-1\}$$

where the $\{x_{ik}\}$ and $\{b_{ik}\}$ are fixed- and random-effects predictors respectively, the $\{\beta_k\}$ and $\{b_k\}$ are fixed- and random-effects regression parameters respectively, and α_i is a THRESHOLD PARAMETER for the boundary between the j-th and $(j+1)$-th response categories. (The threshold parameters play a role analogous to the intercept in an ordinary mixed logit model.) The linear predictors are related to cumulative probabilities $\{\gamma_{ij}\}$ through the inverse logit function:

$$\gamma_{ij} = P(Y_i \le j) = \frac{e^{\eta_{ij}}}{1 + e^{\eta_{ij}}}$$

The probability of datum i having response category j is thus

$$P(Y_i = j) = \begin{cases} \gamma_{i1} & j = 1 \\ \gamma_{ij} - \gamma_{ij-1} & j \in \{2, \ldots, N-1\} \\ 1 - \gamma_{ij-1} & j = N \end{cases}$$

That is, the $\{\gamma_{ij}\}$ carve up the interval $[0, 1]$ into a set of category response probabilities, as illustrated in Figure 8. In mixed-effects cumulative logit models, the random-effects regression parameters are assumed to be drawn from some multivariate normal distribution, the covariance matrix of which is estimated jointly along with the threshold parameters and fixed-effects regression parameters (here, via Laplace-approximated maximum likelihood).

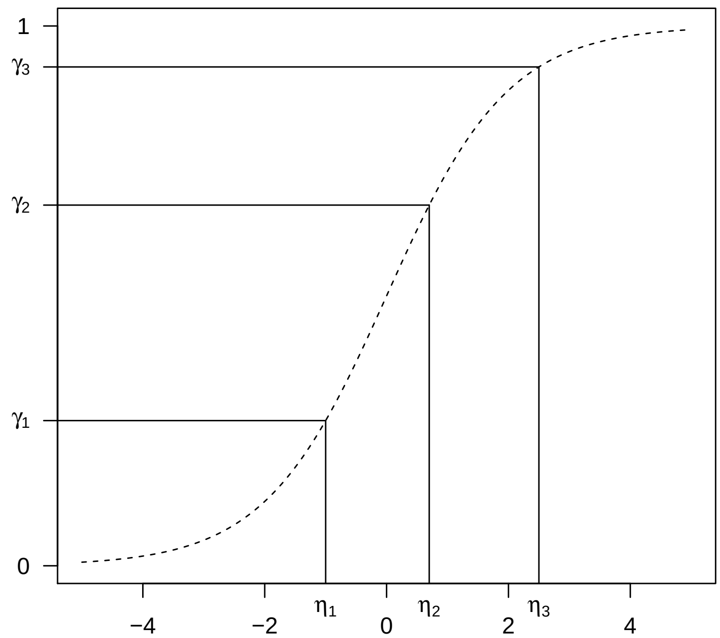

Figure 8: Cumulative logit models. The $N-1$ linear predictors $\{\eta_j\}$ induce a set of N multinomial response category probabilities through the inverse logit transform.

Acknowledgments

We would like to thank David Barner, Victor Ferreira, Ray Jackendoff, Titus von der Malsburg, Emily Morgan, Jesse Snedeker, and Heike Wiese for helpful discussion, and Suhas Arehalli for assistance in running the studies. This work was supported by NSF Grant IIS-0953870, NIH Grant HD065829, and an Alfred P. Sloan Fellowship to Roger Levy, as well as by an German Academic Exchange Service (DAAD) postdoctoral scholarship to Eva Wittenberg.