Semantic Combinatorial Processes in Argument Structure: Evidence from Light-Verbs

MARIA M. PIÑANGO¹, JENNIFER MACK¹ & RAY JACKENDOFF²
¹Yale University & ²Tufts University

1. Introduction
Any theory of how language is internally organized and how it interacts with other mental capacities must address the fundamental question of how syntactic and lexico-semantic information interact at one central linguistic compositional level, the sentence level. With this general objective in mind, we examine “light-verbs”, so called because the main thrust of the semantic relations of the predicate that they denote is found not in the predicate itself, but in the argument structure of the syntactic object that such a predicate licenses. For instance, in the sentence Sue made a dash across the road, the character of the event expressed by make a dash is determined not so much by the main predicate make as by the object DP a dash. This situation defines the verb make as a light-verb, and the combination of the verb and the argument structure of its object as a “light-verb construction”. Their behavior therefore lies at the interface between syntactic and lexico-semantic representation (e.g., Jespersen 1954, Jackendoff 1974, Cattell 1984, Grimshaw and Mester 1988, Baker 1989, Jun 2003, Culicover and Jackendoff 2005).

Here, we seek to chart the time-course of interpretation of light-verbs, and through doing so, compare two possible ways it might take place. The first is a lexical approach, according to which the light-verb interpretation is fully listed as an idiom in the lexicon (i.e., a case of logical polysemy). The second is a compositional approach, according to which light-verb constructions are built “in real time”, as it were; that is, the resulting interpretation of the light-verb construction arises out of the local syntactic and semantic context of the predicate and its object. We hypothesize that if the construction of light-verb interpretation is lexicalized in nature, it should be observable during real-time comprehension as soon as it is licensed. However, if, as the compositional approach suggests, the light-verb construction’s interpretation is built online from the argument
structures of the light-verb and the nominal, it should be observable only some time after having been licensed, in the form of computational cost.

Determining where the light-verb construction lies in this theoretical divide can thus shed light on the larger question of the nature of argument structure, and ultimately on the nature of the connection between syntactic and semantic representation.

The paper is organized as follows: First, we present a more detailed description of light-verbs and the light-verb construction. Second, we discuss the processing considerations that lead to the specific hypothesis pursued here. Third, we present a corpus analysis and the results of two real-time comprehension experiments that test our hypothesis. We conclude with a discussion of where our findings fit in the representation of the syntax-semantics correspondence.

2. The Representation of Light-Verbs: Argument Sharing

Consider the following contrast

1) The man took a cup
2) The man took a walk

Both sentences show the same basic phrase structure representation: [DP [VP V DP]]. The way the meanings of their parts are combined, however, is different. In sentence (1), the interpretation arises straightforwardly from the meaning of the lexical items involved: take licenses two syntactic arguments which map directly onto the semantic roles of <agent, theme> in a one-to-one correspondence. In sentence (2), take again licenses two syntactic arguments. The subject is still <agent>, but the object is not a <theme>. Rather, the semantic argument structure of the complement (a) walk, a deverbal noun, has been allowed to activate the argument structure of its source verb to walk, <agent>; this semantic role is identified with or fused with or shared with the agent role of the main verb, thus giving rise to the attested interpretation THE MAN WALKED.

Thus the interpretation of (2) is strictly compositional: the interpretation of the light-verb+object composite must be viewed as the combination of the argument structure of the main predicate with the argument structure of the predicate’s object. But crucially, this interpretation has come at a cost: the semantic roles activated and the syntactic arguments onto which they are mapped are no longer in the canonical one-to-one correspondence seen in (1). We take this observation to be necessary for any kind of account of light-verb interpretation, and we use the label “argument sharing” to denote this process.

Summarizing, argument sharing refers to the process whereby the semantic system combines at least two sets of arguments by matching as best it can their independent properties. So, if the usual use of take shown in (1) has the argument structure (3), where syntactic and semantic arguments match, then argument sharing in take a walk might be schematized as in (4). (Note that the light-verb stipulates the syntactic role object without a corresponding thematic role, and the
nominal stipulates thematic roles without corresponding syntactic roles. See Culicover and Jackendoff 2005 for a related account.)

3) take <subject=agent₁, object=theme₂>
4) take <subject=agent₁, object₂>
   + walk₂ <agent₃>
   → [take a walk]ᵥp <subject=agent₁,₃>

For a more complex case, we may contrast the canonical use of get in (5) (John got a book from Bill) with the light-verb use in (6) (John got an order from Bill to finish the proposal).

5) get <subject=goal₁, object=theme₂, from NP=source₃>
6) get <subject=goal₁, object₂, from NP=source₃>
   + order₂ <agent/source₄, goal₅, action₆>
   → [get an order]ᵥp <subject=goal₁,₅, object₂, agent/source₄, action₆>

Two main implications follow from this approach to light-verb interpretation:
1) the absence of the expected one-to-one correspondence between semantic representation and its corresponding morphosyntactic structure need not always be a cause of ungrammaticality; 2) the conditions on argument sharing are formal (i.e., capturable in terms of a lexico-semantic structure), not necessarily the result of lexico-conceptual biases created over experience (i.e., world-knowledge).

From a real-time comprehension perspective, argument sharing describes light-verb interpretation in terms of a potentially measurable process with predictable parsing-based consequences. If argument sharing takes place in real-time, its implementation should be more costly to the processing system, as it entails a step above and beyond that which must be carried out in the interpretation of the non-light counterparts.

However, before this last implication can be explored further, we must check on the frequency distribution of the light-verb construction. For if light-verbs take longer to process, it might just be because they are of lower frequency in use. We need to defuse this possible objection.

3. Dissociating “idiomatic” from “compositional” processes: Corpus Analysis

If argument sharing is compositional, it is built during real-time sentence formation. From this perspective, it is only a more complex instantiation of normal semantic role assignment. Alternatively, light-verb constructions might be stored in the lexicon, with composite argument structures of the sort in the lower lines of (4) and (6). On this view, argument sharing is noncompositional in real time: it is a fully lexicalized (i.e., idiomatic) “re-arrangement” of the argument structure of a predicate+argument composite. In this case, there should
be no extra processing cost associated with light-verb constructions beyond normal lexical access for idioms.

In order to tease these two potential accounts apart, we must examine the possibility that if the prediction of extra processing cost is borne out, it is not for the irrelevant reason that there is a basic sense-frequency difference between light-verb constructions and non-light verb counterparts. To this end, we carried out a corpus analysis of light *senses* using the Brown Corpus of Written American English (Francis and Kucera 1964, Hofland et al. 1999). The search was carried out using the following criteria: 1) all verbal forms of a set of potential light-verbs had to be included, 2) each occurrence had to be analyzed within its local context (approximately 15-20 words on either side), 3) light senses counted could only be those that had a non-light counterpart.

These criteria yielded the following sense classification for the set of verbs chosen:

- **Light** senses, containing argument sharing between the verb and its object, e.g., "She changed ends and *took a swipe* at John ..."
- **Dark** senses like (1) and (3), containing a verb with its expected argument structure but no argument sharing, e.g., "I *took the pint bottle* from my pocket and ..."
- **Other category** senses- which include alternative senses of the verb as in: "It *took a tragedy* to bring things to a head" (where *take* = REQUIRE)
- Clearly idiomatic (i.e. non-compositional) uses such as "He *took heart*.

The results for all verbs examined (a total of 11) reveal that the light sense is used more frequently than the dark counterpart [$\text{Mean}_{\text{Light}} = 223$ (261), $\text{Mean}_{\text{Dark}} = 165$ (164), $t_{(10)} = .62$, $p=.27$]. These findings thus suggest that even if there is the hypothesized extra processing suggested by a compositional view of argument sharing, the light sense of a predicate occurs with at least as great a frequency as the dark sense counterpart. Consequently, any observation of increased processing cost during comprehension cannot be attributed to a sense-frequency artifact. This clears the way for our experimental question regarding the time-course of argument sharing, to which we now turn.

4. **The processing of light-verbs: Charting the time-course of the implementation of argument structure**

To formulate our predictions for the composition of argument sharing, we are guided by well-supported notions about the comprehension system. The comprehension process is incremental, and semantic composition that is not directly supported by syntax can be detected in the form of delayed increased processing load. Such a load is not observed in syntactically transparent circumstances, that is, in circumstances where semantic representation has a direct morphosyntactic correlate.
Initial reported evidence strongly supports this hypothesis (Piñango et al. 1999, Todorova et al. 2000, McElree et al. 2001, Piñango et al. 2005, Piñango et al. in press). These reports show that certain types of enriched composition (e.g., aspectual, complement and argument coercion) can be isolated during the course of normal comprehension; this process is detected as a “computational cost” to the processing system.

The basis of the processing approach is as follows: the temporal location of this increase in load is 250-300 msec after a lexical semantic combinatorial operation has been licensed. This observation fits in well with other evidence concerning the time course of combinatorial semantic processing (e.g., Swinney and Smith 1994). Although information concerning potential argument structure configurations seems to be available soon after hearing the verb in a sentence (e.g., Shapiro et al. 1987, 1989), the deployment of such information in thematic role assignment appears to follow a slower-developing, longer-lasting real-time course than do syntactic operations. For instance, in a cross-modal task, Boland (1996) reports this pattern: syntactic anomaly influences naming response time significantly earlier than does meaning-based anomaly. McElree & Griffith (1995), using a judgment task, similarly report violations of thematic role structure to be detected later than syntactic violations in the course of sentence processing. Even though these previous studies address mainly lexical as opposed to compositional semantics, and even though they capitalize on the detection of violations, they point to a clear and detectable difference between syntactic and semantic operations as comprehension unfolds, thus validating the present approach. This means that, if the compositional approach is correct, light-verb interpretation (in the form of argument sharing) should be detected during sentence processing – as delayed increased processing load.

4.1 Methods

Design and Procedure. Our current experimental design makes use of the dual task interference paradigm. In this task, sentences are presented auditorily to subjects. At a certain point in each sentence, a letter string (probe), totally unrelated to the sentence, is presented visually. At this moment, the subject is required to make a lexical decision for that probe i.e., to decide whether or not the letter string forms a word in English. The dependent measure is the reaction time (RT) to this decision. The key assumption underlying the task is that a hypothesized processing load involved in the comprehension of one of the conditions will “interfere” with the processes associated with executing the lexical decision task. This interference will be revealed in a higher RT for that condition over its counterparts. In the present case, the factor hypothesized to cause higher processing load is the presence of argument sharing.

For all three conditions, two positions are probed: right at the end of the matrix object DP, which we refer to as the Licensing Condition (LC), and 300 msecs after the end of the matrix object (LC+300). The RT to the lexical decision
is taken to be an indicator of the processing load at the post-object position of the light condition as opposed to the dark and heavy counterparts.

**Stimuli.** Twenty-five triads of experimental sentences were constructed. Each triad contained a light condition such as (7), where a verb and its nominal object must share their argument structures; a dark condition such as (8), in which the same verb occurs but with a canonical argument structure; and a heavy condition such as (9), in which the same nominal occurs, but in a canonical argument structure construction with no argument sharing.

7) *Mr. Olson gave an order^ last night to the produce guy* (light)
8) *Mr. Olson gave an orange^ last night to the produce guy* (dark)
9) *Mr. Olson typed an order^ last night for the produce guy* (heavy)

Verbs were selected on the basis of the corpus analysis, yielding a set of triads where the frequency of the light sense is systematically higher [Mean\textsubscript{Light} = 263.2 (275.1), Mean\textsubscript{Dark} = 192.2 (171.5), \(t(24) = .65\), \(p = .26\)]. Verbs were also matched for raw frequency and length (in terms of both letters and syllables).

Sentential objects (e.g., order/orange) were also matched for frequency and length (letters and syllables), giving always an advantage to the object of the light condition. This was done so that if there were any processing facilitation due to frequency, it would give the advantage to the sentences with verbs requiring argument sharing, thus blunting the effect we hypothesized – higher processing load for the light condition sentences. Objects were all one or two syllables long. Whenever possible, they were also paired for phonological content of the first syllable to ensure that the post-object position (Licensing Condition) would be as close to the “real” LC as possible (i.e., to align the point during the sentence where the probe appears with the point in comprehension of the object when the content of the lexical item is fully retrieved).

On the idea that formulaic light-verb constructions may in fact be stored in the lexicon, and this would predict a lower RT than the dark/heavy conditions, throughout the creation of the light condition, we sought to avoid idiomatic uses of the “light” constructions.

Twenty-five triads of visual probes, to be aligned with the 25 pairs of experimental sentences, were constructed. The probes were selected so as not to be related to the content of the sentence in any meaningful way. To ensure interference, probes were constructed with a variety of manner of articulation features. Probes were also matched for frequency and number of letters. Probes were left on the screen for 400 msecs.

An additional 225 sentences and probes were constructed as foils. One hundred and twenty-five sentences were assigned nonword probes constructed to conform to orthographic and phonotactic rules of English, and the remaining seventy-five sentences were assigned probes made up of legitimate words of the
English language. Probe position in all filler sentences was randomly varied to lessen expectations on the part of the subjects.

Two final scripts containing a total of 300 sentences were constructed, each encoding a counterbalanced experimental sentence ordering. All 75 experimental sentences were matched to word probes. Of the 225 filler sentences, 75 were matched to word probes, and 150 were matched to nonword probes. Within each of the scripts, the experimental sentences were set up in a quasi-random order, with the constraint that the light sentence of the pair should precede the dark/heavy counterpart about half of the time and the dark/heavy condition precede its light counterpart the other half. Probes of filler sentences were ordered in such a way that no more than three consecutive sentences had a nonword probe.

Construction of the final script (containing both experimental and filler sentences) was guided by the constraint that there be at least one filler between any two experimental sentences and that there be three filler sentences at the beginning of the script, always avoiding any type of block distribution.

All sentences were digitized at 22,000 samples per second. Probes were triggered using TEMPO (Motta, Rizzo, Swinney and Piñango, 2000-2004), software developed in our lab for the implementation of cross-modal techniques.

For the light vs. dark contrast, two experimental lists of probe words were created to use with these sentences. For each experimental position (LC vs. LC+300), half of the subjects were presented with one pair of probes and sentences, and the other half were presented with the same pair of sentences but with the probes exchanged. The intention behind this was that any potential facilitation caused by an unforeseen probe-sentence connection would be detectable by a post hoc interaction analysis between probes and sentences.

Sentence-probe pairs were kept constant in the heavy condition in order to minimize subjects’ exposure to the whole set of experimental sentences. We reasoned that if any effect observed between the light and heavy conditions were to arise from the probe itself, such an effect should be observed in both LC and LC+300 positions. Finally, probe frequency was controlled in the light/heavy contrast. However, any difference in frequency was made to systematically favor the light condition (i.e., H< L&D), so as to make any accumulated facilitation effect from probe frequency work against our predictions.

Subjects. A total of forty subjects (twenty subjects per position) were tested in this experimental series. All were monolingual speakers of English with normal or corrected-to-normal vision and hearing (by their own report).

### 4.2 Predictions

Our hypothesis that argument sharing is compositional and semantic in nature makes two specific predictions: 1) the compositional nature of argument sharing predicts that it should be visible as processing cost, as it demands computations above and beyond those required in the dark and heavy counterparts; and 2) the
semantic nature of this process predicts that the cost of argument sharing should not be visible until some time after licensing; that is, at the LC+300 position. This is so because semantic composition is built at least partly on the basis of syntactic composition, so the results of semantic compositionality are expected to be observable in a protracted manner with respect to the morphosyntactic representation on which it is built (e.g., McElree and Griffith, 1995, Piñango et al. 1999, 2005). In summary, our hypothesis about light-verbs in combination with well-known observations about the processing of semantic composition predict that the cost of argument sharing will be reflected in a statistically significant higher RT only at the LC+300 position to the lexical decision for the light condition over the dark and heavy conditions.

4.3 Results
Subjects responded at ceiling levels (with a success rate of 89.5%) to the comprehension questions, indicating that they consistently attended to and understood the experimental sentences.

There were no group by condition interaction effects observed neither at LC [F(1,36) =.000,  p=.998] nor at LC+300 [F(1,36)=.001,  p=.979] positions. This indicates that any potential factor that could have facilitated or interfered in the processing of the experimental sentences was successfully controlled in our design.

There were two criteria for data exclusion: 1) incorrect or non-existent responses – when a subject answered “no” to an experimental word probe, or failed to respond at all, and 2), delayed responses – when the reaction time for a given item fell outside three standard deviations from the subject’s mean RT for that condition. These criteria resulted in the exclusion of 7.7% of the total data set, distributed equally across the three conditions.

At LC, a subjects-based paired t-test indicated that reaction times were significantly higher in the dark condition than in the light condition: Mean_{Dark} = 634.43 ms (82.34 ms) vs. Mean_{Light} = 619.15 ms (82.84 ms) [t_{(19)} = -2.01,  p<.03, one-tailed]. This effect was not predicted, yet not completely unexpected given the frequency difference between object DP’s in the light vs. dark conditions: object DP’s in the dark condition were systematically less frequent – and therefore presumably more costly to process – than their light and heavy counterparts. In contrast, and consistent with our predictions, no difference was observed between the light and heavy conditions at this position: Mean_{Heavy} = 618.40 ms (86.38 ms) vs. Mean_{Light} = 619.15 ms (82.84 ms) [t_{(19)} = .112, p=.456 (one-tailed)].

Moving on to the LC+300 position, RTs for the light condition were numerically higher than RTs for the dark condition: Mean_{Light} = 664.44 ms (115.43 ms) > Mean_{Dark} = 659.04 ms (112.88 ms), [t_{(19)} = .634, p=.267 (one-tailed)]. Yet, this trend did not reach significance (p >.05). Given the light << dark effect observed at the LC position, we take this trend to be indicative of a reversal in computational cost in the predicted direction, yet, one which did not
have enough time to develop due to the previous frequency effect it had to overcome.

Focusing on the light vs. heavy contrast, a paired t-test (with subjects as the random variable) revealed a statistically significant difference (light >> heavy) at LC+300: Mean\textsubscript{Light} = 664.44 ms (115.43 ms) > Mean\textsubscript{Heavy} = 645.51 ms (104.2ms), [t\textsubscript{19}]=2.46, p=.012, one-tailed. This effect was also obtained in the items-based analysis: Mean\textsubscript{Light} = 666.62 ms (43.30 ms) > Mean\textsubscript{Heavy} = 647.30 ms (40.14ms), [t\textsubscript{24}]=1.718, p=.049, one-tailed. This is consistent with our original hypothesis that argument sharing involves additional semantic composition.

5. Discussion and Conclusions

The representation of argument sharing reveals that it is rooted in syntax and semantics. That is, it occurs in a syntactic compositional process, yet, it is triggered by the presence of a semantic roles- syntactic constituent mismatch. At least in English, it is morpho-syntactically opaque. It does not introduce new content; rather it “recycles” what is already there from the argument structures available (i.e., no new pragmatic or semantic roles that do not already exist in the lexico-semantic representation are added to the sentential interpretation).

The processing of argument sharing further enriches our understanding of it by showing that it is carried out at a computational cost; an effect most clearly observed in the light vs. heavy contrast, which emerges 300msecs after licensing. In line with previous findings on enriched composition, we take this computational cost to be a reflection of a lack of overt morphosyntactic support.

We interpret its protracted time-course to reflect its purely semantic nature (even though its implementation takes place on the basis of syntactic composition).

This is an interesting outcome for linguistic theory. All descriptions of argument structure agree that it must reflect the number of possible syntactic constituents (subject, object, etc.) that have a corresponding lexico-semantic instantiation (agent, theme, etc.). However, current models of linguistic representation differ as to where the locus of this correspondence lies – as an intrinsic component of syntactic structure, as captured in theta theory (e.g., Chomsky 1981, 1995) or as a syntax-semantic interface component as captured in a variety of theoretical proposals such as lexical conceptual structure (e.g., Jackendoff 1990), argument (a-)structure (e.g., Bresnan 2001), and semantic macrorole structure (e.g., Van Valin and LaPolla 1997). Our present results are consistent with an interface approach to argument sharing, and consequently with an approach to semantic role instantiation that goes beyond syntactic structure.

Finally, in terms of the architecture of language, the results are also consistent with the view, implicit in some, but not all, models of language representation that the semantic system can carry out compositional processes of its own, independent of the combinatorial powers of syntax.
References


Combinatorial Processes in Argument Structure


Maria Mercedes Piñango
Yale University
Department of Linguistics
P.O. Box 208366
New Haven, CT 06520
maries.pinango@yale.edu

Jennifer Mack
Yale University
Department of Linguistics
P.O. Box 208366
New Haven, CT 06520
jennifer.mack@yale.edu

Ray Jackendoff
Tufts University
Department of Philosophy
Medford, MA 02155
ray.jackendoff@tufts.edu