Volume xx, Issue xx xx xx

Productivity and Argument Sharing in Hindi light verb constructions

Published by CSLI Publications

Contents

Productivity and Argument sharing in Hindi light verb constructions

Ashwini Vaidya & Eva Wittenberg, *IIT Delhi; University of California, San Diego*

Received Date; Revised Date

Abstract

Light verb constructions (e.g. *give a sigh, take a walk*) are a linguistic puzzle, as they consist of two predicating elements in a monoclausal structure. In the theoretical literature, there has been much interest in the linguistic analysis of such constructions across a range of grammatical frameworks. One such proposal is event co-composition, where the argument structures of noun and light verb merge, resulting in a composite argument structure, which has been claimed to be the source of increased processing costs in English and German. In contrast to these languages, in Hindi a larger proportion of the predicates are light verb constructions. Hence, we may ask whether a Hindi speaker's experience with light verb constructions allows them to go through the same co-composition operation faster than a speaker of English. Our results show that Hindi speakers are adept at the process of using light verb constructions to 'verbalize' predicates, more so than speakers of Germanic languages. We argue that these data provide evidence for a case of specific linguistic experiences shaping cognition: cost disappears with practice.

1 Introduction

One fact that all languages have famously in common is the ability to convey any category of meaning – things, ideas, events, or states. But what varies widely across the globe is how exactly each language packages meaning into a syntactic structure. In this paper, we explore how the prevalence of such a packaging strategy will lead to different observable behaviors in the speakers of a language.

Our main concern is a predicational strategy where a verb such as *do, make*, or *give* combines with an event noun, such as *a jump* or *a call* to form a phrasal structure with a single meaning (*jumping, calling*). Such constructions are widespread and can be found in languages as diverse as Hindi, Persian, and English. They belong to the class of complex verb constructions; in this paper, we focus particularly on light verb constructions (Jespersen 1965).

Light verb constructions open up an interesting perspective on how we conceive of the interaction between the usage patterns, grammatical allowances, and grammatical possibility spaces within a particular language on the one hand, and the fundamentals of the human mind on the other hand: While cognitive mechanisms, concepts in the mind, and expressive needs do not vary significantly across the world, each language has its own grammatical preferences, combinatoric possibilities, and frequency biases. For instance, say you want to ask your child to hug you. In English, the language-wide preferred way to encode an action concept like 'hug' is via a simple verb ('to hug'). However,

JSAL Volume xx, Issue xx, xx xx.

you can also use a light verb construction ('to give a hug'). Crucially, the language-wide preference for simple verbs and the frequency of specific constructions can be independent: While overall, English has few complex predicates like 'to give a hug' (low language-wide systemic frequency), the ones that it has are highly frequent (high token frequency).

In other languages, the picture is different: In Hindi for instance, there is a language-wide preference to encode actions as complex verbs, thanks to a very productive light verb construction schema. In either language, comprehenders need to construct the path from a linguistic structure to an action concept, relying on the same mental architecture across languages.

Here, we ask whether language-wide preferences, not only token frequency, interact with this comprehension process: Light verb constructions have been analyzed as a general process of event *co-composition*, one example of constructions that compose predicative meaning from both an event noun as well as from a verb (Ahmed et al. 2012). Given that this predicational strategy is present across many languages and language families, but also given each language's systemic preferences for one predicational strategy over the other, we can ask whether a speaker's ample experience with light verb constructions allows her to go through the same cognitive event co-composition operation faster than a speaker of a language in which light verb constructions are less productive.[1]

There has been psycholinguistic research on the processing of light verb constructions, which has found evidence for a cost of co-composition (Piñango et al. 2006, Wittenberg et al. 2014, Wittenberg and Piñango 2011). However, the languages studied in these experiments (English and German) use light verb constructions overall relatively infrequently. In contrast, this paper asks whether the process of event co-composition is observable in Hindi, a language that uses complex verbs for nearly a quarter of its predicates. That is, if we take the nature of the cognitive process of event composition as identical across languages (and speakers' minds), we would still expect the overall frequency and productivity of co-composition in Hindi, and cognitive constraints, such as working memory (Norcliffe et al. 2015), to interact with grammatical processes.

Our paper is organized as follows: We begin with a theoretical description of the event composition process. We then examine evidence for this process in the form of existing psycholinguistic studies, before turning to a discussion of the light verb construction in Hindi. Following this, we describe four experiments that compare how light verb constructions and their non-light counterparts are processed, using two different experimental techniques. We conclude with a summary and discussion of the results.

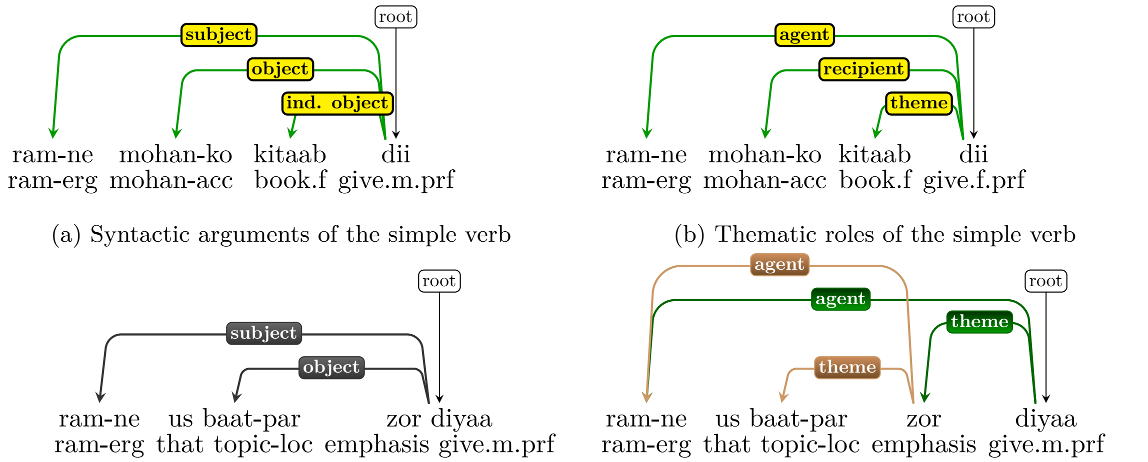
1.1 A model of co-composition

Many theoretical studies across grammatical frameworks have proposed analyses of light verb constructions, because they are challenging for the interaction of semantic and syntactic information in language (Sag et al. 2002, Culicover and Jackendoff 2005, Wittenberg 2016). Specifically, in most of the literature, the theoretical questions are related to the formal representation of compositionality: If a light verb and a noun together form a predicate, then how does the syntactic and semantic representation of both these elements result in the particular syntactic and semantic properties of the light verb construction?

Usually, a verb (for instance, *to describe*) denotes the meaning of an event (in that case, someone uttering something) and, its object or objects (for instance, *a dance*) will fill in argument slots: *describe a dance* denotes a recount of a dance. In light verb constructions (for instance, *do a dance*), the verb (*do*) does not supply the event type – we know that we are talking about a dancing event because the predicative meaning is supplied by the syntactic object (*a dance*).

One account that effectively handles this problem argues that the argument structures, demanded by the lexical semantics of both noun and light verb, overlap with each other, and a 'shared', or 'composite' argument structure, emerges in the monoclausal light verb construction (Mohanan 1997,

[1] Note that this is neither a Whorfian question – there is no claim that there are language-specific influences on perception – nor is it a question about token frequency of a particular construction. Rather, we ask how language-wide structural preferences interact with cognitive processes.



(a) Syntactic arguments of the simple verb

(b) Thematic roles of the simple verb

(c) Syntactic arguments of the light verb

(d) Composite thematic roles of noun and light verb

FIGURE 1: A simple verb with its syntactic arguments is shown in Figure (1a). The same simple verb maps the syntactic arguments to the semantic ones in (1b). The light verb has two surface syntactic arguments in (1c). But these emerge after two sets of thematic roles from noun and light verb combine in Figure (1d), resulting in a composite argument structure.

Alsina et al. 1997, Ahmed et al. 2012, Durie 1988, Piñango et al. 2006). For instance, the Hindi Ex. (1) consists of the ditransitive verb *give* 'de' with three arguments, *Raam*, *Mohan* and *kitaab* 'book'. In this case, three syntactic arguments, subject, indirect object and object, are mapped directly to the set of thematic roles for the verb, agent, recipient and theme (Figures (1a) and (1b)).

(1) raam=ne mohan=ko kitab **d-ii**
Ram.M.Sg=Erg Mohan.M.Sg=Dat book.F.Sg give-Perf.F.Sg
'Ram gave Mohan a book'

(2) raam=ne us baat=par **zor** **di-yaa**
Ram.M.Sg=Erg that topic=loc pressure.M.Sg give-Perf.M.Sg
'Ram put an emphasis on that topic'

On the other hand, Ex. (2) has the light verb *de*, which takes two syntactic arguments: *Ram* and *us baat* (Figure (1c)). These emerge only after the two sets of thematic roles from the noun and light verb are combined. Both noun and light verb contribute towards the argument structure of the sentence. The light verb *de* licenses the nominal predicate argument *zor* and the agent *Raam*. The noun *zor* also licenses the agent *Raam* and the theme *us baat par* (see Figure (1d)).

In language comprehension, the agent argument *Raam* must be identified as common to both predicates *zor* and *de*. This is sometimes described as the *argument identification step* (Davison 2005). In Davison's (2005) model, following argument identification, a semantic argument merger step will take place such that a composite, but monoclausal structure is formed. This interaction between the syntactic arguments on the surface and the two sets of semantic arguments is characteristic of light verb constructions, and it has been defined as event co-composition: a process where "two semantically predicative elements jointly determine the structure of a single syntactic clause"(Mohanan 1997, p. 432).

Light verb constructions have additional properties that are crucial to the event co-composition process. Across languages, light verbs consist of a small class of high-frequency, general-purpose verbs that are form-identical with their non-light counterpart, e.g. *make, do, give, take* etc., a fact that has led Butt (2010) to consider the light verb as a unique category that shares a lexical entry with its non-light form. That means that at one point during comprehension, the listener or reader needs to interpret the light verb as light rather than non-light, in order for the the event composition process to succeed. For example, in (1), we have the verb *de* 'give' in its non-light form, but in (2), it is part of a light verb construction.

An effect of this form-identity with full verbs is that when a comprehender resolves a verb as light, the predicating noun provides the eventive meaning, but the verb still supplies additional aspectual or agentive information about the event and its structure (Butt et al. 2008, Wittenberg et al. 2017). This also has implications for the structure and nature of the semantic arguments expressed in a light verb construction. For instance, while the non-light use of *give* (e.g *give an orange to someone*) has three semantic roles (Source, Theme, and Goal), the light use (e.g *give a kiss to someone*) has only two (Agent and Patient), with the predicating noun *kiss* fusing its agent role with *give*. These syntactic and semantic interactions with the predicating noun result in the structure of a complex predicate. In the next section, we review the results from the psycholinguistic literature that have shed light on how the process of light verb construction comprehension unfolds.

1.2 Measuring event co-composition

In light verb constructions, the challenge for comprehenders is to understand that the verb, usually the only projector of sentential argument roles, is sharing this power with the event nominal. This means that the event co-composition process lies at the interface between syntactic and lexico-semantic representation. During real-time interpretation of such constructions, how do comprehenders resolve the mismatch between the syntactic and semantic arguments in the clause?

The literature offers several psycholinguistic studies on light verb construction comprehension in English and German. These studies have focused on collecting behavioural and electro-physiological data at the point when the verb is read or heard (in Subject-Object-Verb or Object-Verb-Subject structures) or at the noun (in the case of Subject-Verb-Object structures), based on the theoretical prediction that the composite argument structure of noun and light verb must be 'resolved' after both verb and noun are processed, and the co-composition process would be observable as a behavioural or electro-physiological signal.

Briem et al. (2009) carried out three experiments to study how light verbs are processed in German. They contrasted light verbs like *geben* ("give") with non-light verbs like *erwarten*("expect") either in contrast with a pseudo-word, by themselves, or within a sentential context. Briem et al. (2009)'s study used MEG in order to demonstrate that light verbs (e.g.*give*) when presented by themselves or in comparison with pseudo-words showed less cortical activity as compared to non-light verbs (e.g.*expect*). The authors interpreted this as a result of lexically underspecified features in the light verb. In the third experiment, when presented in a minimal sentential context using object-verb-subject structure, a verb like *give* had to get resolved as either light or non-light depending upon the presence of the noun, e.g. *a kiss gives he* vs. *a book gives he*. Here, the pattern followed that of the previous two experiments; a non-light context *give a book* resulted in greater left-temporal activation as compared to light. Briem et al. (2009) interpreted these results as evidence for distinct brain processing areas for distinct categories of verbs. However, the authors did not measure activity *after* the verb, where subsequent behavioural studies have found differences between light and non-light conditions (for further discussion, see (Wittenberg et al. 2014)). We now turn to some of these experiments which report such later effects.

A handful of behavioural experiments have studied late reflections of computational costs associated with light verb construction processing, motivated by the mismatch between the syntactic argument structure and the semantic roles defined by the construction (see section 1.1). This mismatch was predicted to surface as computational cost *after* the construction has been comprehended.

Piñango et al. (2006) predicted to show a cost of processing light verb constructions at around 250-300ms after the offset of the construction. This prediction was based on studies that had shown that the (re-)assignment of semantic roles results in slower-developing effects that can be detected at a later point during sentence processing (Boland 1997, McElree and Griffith 1995). Based on this idea, Piñango et al. (2006) used an interference paradigm in the form of a cross-modal lexical decision task. Piñango et al. (2006) used three conditions for their study, where the light condition was contrasted with non-light condition, and a third condition ('heavy') containing the same noun, but paired with a non-light verb:

(light) Mr. Olson gave an order last night to the produce guy
(non-light) Mr. Olson gave an orange last night to the produce guy
('heavy') Mr. Olson typed an order last night for the produce guy

Participants listened to sentences containing one of the three conditions as shown above. At a point after the object (*order* or *orange*) was heard, participants had to make a lexical decision on a letter string that flashed on the screen. The reaction time to make a decision was taken as a reflection of the demand placed on working memory by the construction that was just heard: Slower reaction times would reflect higher computational cost.

Piñango et al. (2006) also manipulated the timing of the probe placement: Either the probe was placed immediately at the offset of the object, or 300ms after. In the former case, the non-light condition was significantly slower, an effect that Piñango et al. (2006) attributed to the higher frequency of the light verb construction. However, when the probes were placed 300ms after the noun was heard, this effect was numerically reversed (albeit with no statistically significant difference), and the light condition elicited significantly slower reaction times than the heavy condition. Piñango et al. (2006) concluded that the computational costs of argument sharing become apparent when measured later, i.e after the construction is 'disambiguated' as light.

Wittenberg and Piñango (2011) replicated this study in German. Like Piñango et al. (2006), they also used three conditions (light, non-light, and heavy), but since German can be verb-final (e.g., *a hug give*), the probe was placed either immediately or 300ms after the verb (not noun) was heard. The interference immediately after the verb resulted in no significant differences between the conditions. But again, when the probe was placed 300ms after the verb, listening to light verb constructions while making a lexical decision resulted in significantly slower reaction times than listening to either non-light or heavy constructions. This pattern of results was interpreted by Wittenberg and Piñango (2011) as further support for the hypothesis that the event co-composition process can be measured as a late, more gradually developing effect, after the lexical composition process has taken place.

However, the cost of co-composition could not be replicated in a self-paced reading study in German using the same stimuli (Wittenberg 2013). In this self-paced reading experiment, people read light and non-light constructions at similar speed, with only semantically anomalous constructions being processed slower. We return to the results of this experiment later in the paper.

Yet another study (Wittenberg et al. 2014) used Event-Related Potentials to understand the processing of light verb constructions. This study had three conditions: light (*give a kiss*) and non-light (*give a book*), like the interference-based paradigms, and then an anomalous condition, consisting of a non-felicitous noun-verb pairing (**give a conversation*). This was done mainly to distinguish the processing of complex, but plausible, event structures from implausible ones. The authors found evidence for a late, widely distributed, but frontally focused negativity after the onset of the light verb, compared to the non-light counterpart, and the anomalous condition showed a larger positive effect associated with semantic anomalies. Wittenberg et al. (2014) interpreted their results as reflecting a working memory cost caused by the process of event co-composition, particularly the linking of the two syntactic argument structures that surface in a single monoclausal structure.

In sum, the results of these studies generally point towards two broad themes: first, that the cost of event co-composition can be measured after the verb and noun combination has been processed, using both a cross-modal task and electro-physiological methodologies, but that the effect could not be detected in self-paced reading. Second, this cost appears to be a late effect, developing several hundred milliseconds after the light verb construction has been licensed. Its signature is distinct from semantically implausible constructions, but consistent with the processing of other types of complex events. Taken together, three out of four studies provide evidence for the cost of event co-composition in English and German.

1.3 Light verb constructions in Hindi

Light verb constructions are found across South Asian languages, including Hindi (Masica 1993). Seiss (2009) identifies the following properties that distinguish them from other types of verbal multiwords. First, they are always form-identical with the main or 'full' verbs in the language. Second, they are restricted in their combinatorial possibilities with the predicating noun, that is, every light verb can only combine with a certain kind of noun; and finally, they contribute subtle semantic information in the form of telicity and agentivity, among others (Hook 1974). In describing these construction, we need to acknowledge terminological differences: in the South Asian linguistics literature, light verbs are sometimes subsumed under the term 'complex predicates'. While this term is arguably used more widely, we deliberately use the term *light verb construction*, because in Hindi 'complex predicates' may refer to noun and light verb combinations as well as verb and light verb, or even adjective and light verb combinations. Here, we use 'light verb constructions' to only refer to complex predicates consisting of a predicating noun and a light verb.

In a Hindi light verb construction, a verb combines with another pre-verbal noun, predicate, adjective, adverb, borrowed English verb, or noun (Ahmed et al. 2012). In this paper, we focus only on constructions with nouns, to keep comparability to previous studies. In the sentences below, the verb *de* 'give' is used as simple predicate in Ex. 3, but in Ex. (4), the verb *de* is light and combines with a predicating noun (both examples repeated from the previous section).

(3)	raam=ne	mohan=ko	kitab	d-ii
	Ram.M.Sg=Erg	Mohan.M.Sg=Dat	book.F.Sg	give-Perf.F.Sg

'Ram gave Mohan a book'

(4)	raam=ne	us	baat=par	**zor**	**di-yaa**
	Ram.M.Sg=Erg	that	topic=loc	pressure.M.Sg	give-Perf.M.Sg

'Ram put an emphasis on that topic'

In English, many light verb constructions have a denominal verb counterpart (e.g., *take a walk* can also be expressed by *walk*; (Tu and Roth 2011)). In Hindi as well, some light verbs (like *khoj kar* 'search do') will co-exist with their denominal verbs (*khojnaa* 'to search').

However, the distribution of Hindi light verb constructions differs from English and German because the vast majority of the light verb constructions in Hindi do not have an denominal verb counterpart. While both light and denominal verbs co-exist in English, the formation of denominal verbs in Hindi has ceased to be freely productive (Davison 2005).

Butt (2010) notes that Hindi light verb constructions act as a verbalizers in order to create new predicates and to incorporate borrowed items into the language (e.g. *email kar* 'email do; email'). Light verb constructions are highly productive and are sometimes described as "a preferred way of augmenting the creative potential of the language" (Kachru 2006)[93]. This is reflected in corpora: If English has approximately 7000 simplex verbs, Hindi has only 700 (Vaidya et al. 2013).[2]

1.4 Frequency and co-composition

In all of the experiments that were discussed in section 1.2, the cost of argument structure composition was interpreted as being due to the real-time processing of the light verb construction, because composite argument structures are built 'on the fly'.

An alternative view to this would be that light verb constructions are stored (in the manner of non-compositional idioms) in the lexicon. In order retrieve the right syntax-to-semantics mapping, native speakers would merely detect the construction as light, and retrieve the stored argument structure associated with a given construction. If light verb constructions were stored and retrieved

[2]Based on counts from English PropBank and Hindi PropBank, respectively.

as non-compositional units like this, instead of assembled incrementally and compositionally, what one would predict for real-time processing is that the higher the frequency of a given construction, the faster the recognition; reaction times for light verbs should therefore be faster.

Crucially, in all of the experimental results reported in section 1.2, the token frequency of any given light verb construction was higher than its non-light counterpart. For example, *make*, *have*, or *give* are more likely to occur with a light noun (forming a light verb construction) than with a non-light noun in English or German. That is, the collocational frequencies of light verb constructions such as *give someone a hug* are higher than the collocational frequencies of non-light constructions such as *give someone a book*.

Based on collocational frequency alone, then, one should expect speakers of English and German to be able to process light verb constructions with more ease than their non-light counterparts. However, the results of the psycholinguistic studies do not concord with this prediction (Piñango et al. 2006, Wittenberg and Piñango 2011). In fact, what we see is the reverse: Reaction times are slower despite higher frequency. This pattern of results was interpreted as the cost of co-composition overriding any advantages of collocational frequency, when measured in reaction times at the verb. These results demonstrate that at least in languages like English and German, where the language-wide productivity of light verb constructions is relatively low, the higher collocational frequency of an individual light verb construction does not facilitate processing.

At the same time, as mentioned above, there is a great deal of cross-linguistic variation in how frequently light verb constructions are used. Using a corpus of 50 Wikipedia articles, Vincze et al. (2011) estimated that in English, about 9.5% of the predicates are expressed by light verb constructions. In Hindi, the proportion of light verb constructions is about 37% of roughly 37,600 predicates in the Hindi Treebank (Vaidya et al. 2013). Hindi is not an exception when it comes to language-wide frequency of complex predicates: In a language like Persian for instance, only about 115 simple verbs are commonly used, whereas almost all the rest are light verb constructions (Sadeghi 1993). These numbers highlight the differences in systemic frequency of the light verb construction across languages, and with it, the grammatical productivity of expressing a predicate with a complex verb in a given language.

Thus, in a language like Hindi, where language-wide frequency of the light verb construction is higher than English, we could expect to find that the light verb constructions behave similarly to English, i.e the overall token and language-wide frequencies do not facilitate processing when measured at the verb. Alternatively, we may also find that the overall systemic productivity of the light verb construction results in greater exposure to the type of composite argument structures associated with this construction. This sensitivity towards previously seen argument structures makes them easier to process (Mitchell et al. 1995). If this is the case, then we would expect that the systemic productivity of light verb constructions in Hindi would facilitate processing. Previous exposure to light verb constructions could imply that such argument structures are stored, or that Hindi native speakers are much more efficient at the process of event co-composition itself.

Some cross-linguistic studies support this idea. Structural preferences in different languages will correspond to the frequency with which they appear in those languages. For instance, preferences in relative clause attachment to the head noun in ambiguous sentences seem to differ cross-linguistically. While French and Spanish prefer 'higher' attachment, i.e. to a noun higher in the structure, English and Italian pattern 'lower' (Cuetos et al. 1996). These preferences may be tied to the frequency with which these structures appear across languages (although see Grillo and Costa (2014)'s paper which suggests that other factors may also be involved).

1.5 Frequency and predictability

Token frequency of a particular light verb is context-independent. In comparison, a word or lexical item's predictability depends upon its immediately preceding context, and the particular collocational frequency of a light verb and noun composition can play a role in facilitating its retrieval. Eye tracking studies have shown that the effects of frequency and predictability on reading are distinct and additive in nature (Kennedy et al. 2013). This implies that if a word is both low frequency and unpredictable, it will have greater cost than a word that is high frequency and predictable. There is evidence for the effect of both frequency and predictability on reading times, and Staub (2011) have also shown at these are distinct factors that do not necessarily interact with each other.

In the context of light verb constructions, it is close to impossible to control for the collocational frequency of a noun and light verb (it is almost always likely to be greater than the non-light). But we can control for the *predictability* of both light and non-light constructions, such that they are matched. This will help us tease apart the effect of familiarity or frequent exposure to event co-composition (as a result of language-wide frequency) vs. exposure to the individual noun-light verb combination in its token frequency. Crucially, if both light and non-light constructions are low in predictability, then any facilitation in processing can be attributed to systemic frequency, and not to the individual collocation.

1.6 Experimental Predictions

In the previous sections, we have elaborated on the theoretical motivation for event co-composition, the measurement of this phenomenon using behavioural and electrophysiological paradigms, and its relationship with frequency and predictability. With respect to the processing of Hindi light verb constructions, our experiments ask whether we can replicate the English and German data pattern in Hindi, a language that uses light verbs much more frequently as a predicational strategy than Germanic languages.

If comprehenders across the globe perform co-composition the same way, we should replicate the previous results, with light verb constructions taking longer to process than non-light constructions. However, if the systemic prevalence of light verb constructions in a language (and consequently a greater exposure to those constructions) influences the speed at which comprehenders perform cognitive operations such as co-composition, then we would expect light verb constructions in Hindi to be processed *faster* or equally fast as non-light constructions.

In order to account for the predictability of individual lexical collocations, we control for the predictability of light verb constructions and their non-light counterparts. If light verb constructions are processed faster than non-light constructions under this manipulation, we can conclude that any difference found in previous studies is not due to individual items' predictability, but to adeptness with complex verbs as a predicational strategy.

We test these predictions in four experiments, one of which is a self-paced reading study, and the remaining three use the cross-modal lexical decision task paradigm.

2 Experiments

In this section, we report four experiments on the comprehension of light verb constructions in Hindi, to understand whether the high frequency of complex predicates will lead to different processing patterns from English and German.

2.1 Experiment 1: Self-paced reading

Experiment 1 was designed to ask whether light verb constructions incur a processing cost, compared to their non-light counterparts, in a self-paced reading study. As Hindi is a verb-final language like German, the light verb will likewise appear at the end of the sentence. If we were to find results similar to those found in German, we would expect a difference in the processing of the light condition relative to the non-light condition in the verb region or right thereafter. In addition to both light and non-light conditions, we also include an anomalous control condition that combines a light verb with an incompatible noun (see Table 1), to distinguish effects that are due to semantic implausibility from those that could be a result of event co-composition. A similar control condition was used for German, both in a behavioural as well as in an ERP task (Wittenberg 2013, Wittenberg et al. 2014).

We predict that analagous to Wittenberg's 2013 results, the anomalous condition will trigger longer reaction times compared to light or non-light constructions, because of the semantic in-

Context phrase:	apne samay=ka prabandhan karnaa mushkil | hai isiliye... own.obl time=Gen management.M do.inf difficult be.Pres.sg therefore .. 'It is difficult to manage one's time, therefore ...'
Light/Non-Light/Anomalous:	adhyapak=ne vidyarthi=ko **calender/bhaashan/*silsilaa** diyaa ... teacher=Erg student=Acc calender/speech/*happening give.perf.3.M.Sg 'the teacher gave the student a calender/speech/*happening ...'
Continuation	aur kuch aasaan upaay bhi bataaye and some simple solution.pl also tell.Pl '.. and gave (him) some useful suggestions'

table 1: Example sentence showing all three conditions.

compatibility between noun and light verb in anomalous constructions. For light and non-light constructions, we predict that if the frequency of complex verbs as a predicational strategy influences speed of co-composition, then light verb constructions will be processed faster or equally fast as non-light constructions at the verb and thereafter. But if comprehenders across languages perform co-composition similarly, we would expect longer reading times for light verb constructions, compared to non-light constructions.

2.1.1 Method

Participants read sentences in a masked word-by-word self-paced reading paradigm. We used Ibex Farm for presentation (Drummond 2007). Participants were recruited using Amazon Mechanical Turk. We included a participant screening task that included a series of 8 puzzle questions. Participants were asked to choose between two Hindi sentences, where one was grammatical and the other contained an agreement error. This ensured that the participants were able to make basic grammaticality distinctions in Hindi. This test was introduced before the self-paced reading items were shown as a way to prevent non-Hindi speaking Turkers from participating.

The experiment was preceded by four practice items, followed by 15 experimental items in a Latin square design. We also included 20 fillers, half of which were semantically anomalous. Each experimental and filler item was followed by a comprehension question about the sentence, with two choices (Y/N).

2.1.2 Materials

Fifteen experimental sentences were created for three conditions: light, non-light and anomalous, all using the verb *de* 'give', which can appear both in light and non-light contexts.

Each sentence consisted of a short context phrase, followed by the main sentence ending with the verb *diyaa* and a continuation. Table 1 shows an example sentence across three conditions: light, non-light, and anomalous. All the stimuli sentences were minimal pairs with either a non-light, light, or anomalous noun. In the example shown in Table 1, these are *calendar/speech/*happening* respectively, where the noun **happening* is semantically anomalous in combination with *diyaa*. A list of all experimental items used in this experiment is given in the appendix.

Frequency Norming. As mentioned in section 1.3, Hindi light verb constructions are highly productive. While it is impossible to control the productivity and frequency of a construction in a speaker's language system overall, we can control for individual frequency of a word. Thus, we matched the frequency of pre-verbal nouns across conditions.

To obtain frequency data, we used a corpus consisting of 17 million tokens from BBC Hindi (6.5

million) and the Hindi Wikipedia (10.5 million). This corpus was tokenized and tagged with parts of speech to calculate the frequencies of the nouns as well as the collocational frequencies of the noun and light verb (Reddy and Sharoff 2011). As expected, the collocational frequency of noun and verb was greater for light verb constructions (Mean: 13.62 pairs per million) than non-light constructions (Mean: 1.04 pairs per million); t=3.68, $p = 0.002$ in a two-sample t-test. As expected, anomalous constructions were significantly lower in frequency (Mean: 0.027 words per million) as compared to non-light (t=2.19 p=0.04 in a two-sample t-test). Anomalous were also much lower compared to light verb constructions (t= 4.5, $p < 0.0001$ in a two-sample t-test) (Although the number for anomalous should have been zero, two anomalous nouns *pasand* 'like' and *ghoshanaa* 'declaration' had counts of 8 and 4 respectively, perhaps due to tagging errors in the corpus. However, the anomalous pairs are indeed semantically anomalous).

We were able to control the frequency of the preverbal noun across all three conditions. On average, nouns in the light condition appeared 91.3 times per million, in the non-light condition 92.9 times per million, and in the anomalous condition also 92.9 times per million tokens. There were no significant differences in noun frequency across light and non-light conditions using a two-sample t-test (t=-0.05, p=0.96) or anomalous and non-light conditions (t=-0.0004, p=1).

Acceptability Norming. We also conducted acceptability ratings across all three conditions with 16 native Hindi speakers, who were students of IIT Delhi (11 males, average age: 21.9). Participants were asked to rate sentences on a 7-point scale, ranging from 1-Unacceptable to 7-Acceptable. The average acceptability rating for the light sentences was similar to the non-light (6.22 light, SD=0.85; 5.64 non-light, SD=0.96), while the anomalous sentences had an average rating of 2.97 (SD=1.4). There was no significant difference between the acceptability ratings of light and non-light in a two-sample t-test ($t = 1.74$, $p = 0.09$). There were significant differences in the ratings between non-light and anomalous in a two-sample t-test (t=6.06, p<0.001), which is to be expected.

Cloze probabilities norming. 44 Hindi native speakers, who were students of IIT Delhi (24 males, average age: 22.5), provided sentence continuations for the verb in all three conditions, such as in 8, which had a missing sentence-final verb. The participants were requested to complete the sentence in the most natural way possible.

(8) rohan=ki daadi=ne use promotion milne par badhaii ...
rohan=Gen grandma=Erg him promotion get.Inf on congratulations ...
'On getting promoted Rohan's grandmother (gave) him congratulations'

The verb in the light verb condition was predicted 76% of the time, significantly more often than the verb in the non-light (49.88%) and the anomalous conditions (0.2%) A two-sample t-test showed significant differences in the light and non-light cloze predictions $t = 2.83$, $p = 0.008$). For the light and anomalous cloze predictions as well, a two-sample t-test showed significant differences t=6.68 and p<0.00001. Anomalous and non-light also showed a signficant difference in a two-sample t-test t= 3.83, p<0.0001. Thus, the verb in the light condition was highly predictable, compared to the other two conditions. This is similar to data from German (Wittenberg et al. 2014).

2.1.3 Participants

154 participants completed the experiment on Amazon Mechanical Turk. We selected only those participants who scored above 75% in the Hindi agreement puzzle questions (they had to get at least 6 out of the 8 questions correct). These participants on average had a comprehension score of 88%. This resulted in a total of 101 participants with a mean age of 33.5 years (21 females).

2.1.4 Results

We fit a linear mixed model to log-transformed reaction times with condition as fixed effect; the conditions were treatment-coded with the reference level as non-light, and items and participants were random intercepts (including random slopes for items and participants resulted in non-convergence). The t-values from the linear mixed model were approximated to p-values. The `pnorm` function in R was used to compute the probability density of the region above the obtained t-values. Since this is

Region		**Noun**	**Verb**	**Post-verb-1**	**Post-verb-2**	**Sentence end**
Light vs. Nonlight	t-value	-0.76	0.45	-0.56	-0.92	-0.14
	p-value	0.44	0.65	0.57	0.35	0.89
Anomalous vs. Nonlight	t-value	1.0	**4.35**	**2.73**	0.48	0.75
	p-value	0.32	**<0.001***	**<0.01***	0.63	0.45

table 2: T-values and p-values in the regions after the noun in Experiment 1. The critical region is at the verb. Significant effects in bold.

Region	**Noun**	**Verb**	**Post-verb-1**	**Post-verb-2**	**Sentence end**
Light	753.38	724.82	591.28	580.27	593.32
Non-Light	709.97	676.71	566.36	573.79	569.59
Anomalous	783.61	813.32	637.5	587.79	584.82

table 3: Reading times for comparison between the three conditions for the regions following the noun until sentence end, in milliseconds. The critical region is at the verb, where the difference in the anomalous and non-light condition is significant.

a two-tailed test, the obtained probability value is then multiplied by 2 to give us the approximated p-values. Table 2 gives an overview of the significance pattern for the regions following the noun.

The three conditions did not differ significantly in the regions preceding the noun. At the noun itself, we did not find a difference in reading times between conditions (see Table 3). Reading times at the verb indicate that the verb in the anomalous condition was read significantly slower than the verb in the non-light condition ($t = 4.34$, $p < 0.0001$), but there was no significant difference at the verb between the light and the non-light conditions ($t = 0.45$, $p = 0.65$). The slowdown incurred by the anomalous condition also carried forward to the first postverbal region, where the difference between non-light and anomalous was still significant ($t = 2.72$, $p = 0.006$), but for the light vs. non-light conditions, there were no significant differences in reaction times after the verb was read. For t-values (and their approximated p-values) across all regions in the sentence, please refer to Table 4 in the Appendix. Table 5 in the Appendix also provides the mean reaction times (and SDs) for all regions in the Appendix.

2.1.5 Discussion of Experiment 1

This study compared reading times between light verb constructions, non-light constructions, and anomalous constructions. We did not find any differences in reading times between the light and non-light conditions, although the anomalous condition was read significantly slower than the other two at the verb, and in the region immediately following the verb.

This work is directly comparable to Wittenberg (2013)'s German self-paced reading study, which also included our three conditions (light, non-light, and anomalous), and showed a similar pattern of results as Experiment 1: a slower read anomalous condition and no detectable difference in reading times between light and non-light constructions.

Thus, our Experiment 1 serves as a conceptual replication of the study in German. At the same time, as discussed in Wittenberg (2013) as well, self-paced reading paradigms may not be able to detect more elusive semantic effects. While this method has been shown to reliably detect semantically or syntactically *implausible* constructions (Mitchell 2004), it may be less effective at capturing the plausible but subtle event co-composition processes in light verb constructions.

In section 1.2, we had reviewed the processing of light verb constructions in English and German using interference tasks, specifically the cross-modal lexical decision task (Piñango et al. 2006, Wittenberg and Piñango 2011). Such a paradigm was used to detect the effects of processing costs by placing an additional demand on working memory (e.g. Piñango et al. 2006, Kamienkowski et al.

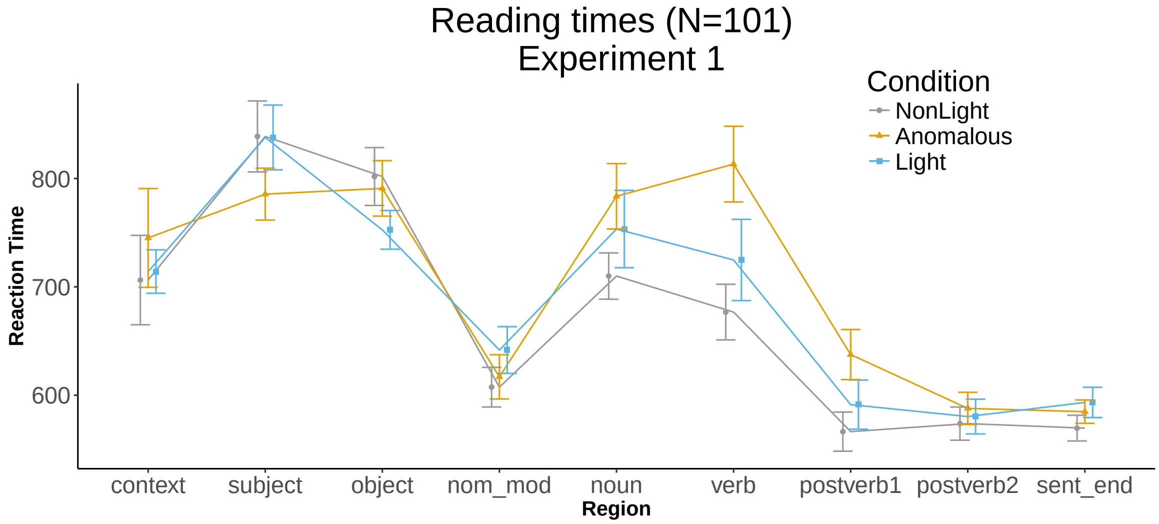


FIGURE 2: Mean reaction times for anomalous, light and non-light conditions for nine regions of interest in the sentence. The error bars show standard error. The critical region was at the verb, where the anomalous condition is significantly different from the non-light condition. Light and non-light are not significantly different at any region in the sentence. The error bars represent standard error.

2011). In an interference paradigm, a deliberate interference with working memory following the verb may slow down the processing of structures which are in fact semantically plausible and grammatical, and may otherwise be processed similar to any other grammatical sentence, but posit an increased demand on working memory due to resolving the mismatch in syntax and semantics, like light verb constructions. In the experiments that follow, we use the cross-modal lexical decision task to test our predictions. Using the cross-modal decision task, we can manipulate the temporal placement of the probe in order to capture these effects.

2.2 Experiment 2

In the previous experiment, we found no reliable differences between reading the light and non-light constructions using a self-paced reading study. In this set of experiments, we decided to investigate the same questions using a different paradigm, particularly to understand whether an enhanced demand on working memory at the verb would capture any fine-grained differences between the conditions.

Hence, this experiment was also designed to ask when differences in processing between light and non-light would be apparent. Just like previous German and English studies (Wittenberg and Piñango 2011, Piñango et al. 2006), we placed the probe immediately after the verb (in this experiment) and 300ms after the verb (in Experiment 3), and measured reaction times to the lexical decision.

2.2.1 Method

Both experiments use a cross-modal lexical decision task paradigm. Participants heard sentences in a light, non-light or anomalous condition. In Experiment 2, a string unrelated to the sentence was visually presented immediately after the verb was heard; in Experiment 3, the probe was placed 300ms after verb offset. Participants had to decide whether the string was a word or a non-word (lexical decision). The sentences were pseudo-randomized in a Latin square design, and the same probe word was used in all three conditions.

Each participant also heard 25 filler sentences, and was asked 20 comprehension questions on both the filler and experimental items. Out of the 25 filler sentences, 15 were semantically anomalous. Thus, each participant heard a total of 40 sentences, of which half were semantically anomalous.

After each sentence, there was a pause of 1500ms and then the next sentence was heard. Each experiment was preceded by a trial session where participants were familiarized with the task. The cross-modal lexical decision task was coded using a browser-based presentation software, jsPsych, version 5.0.3 (de Leeuw 2015), with a custom plugin for the cross-modal lexical decision task paradigm.

2.2.2 Materials

The sentence materials for this experiment were identical to the ones created for Experiment 1. Experimental and filler sentences were recorded by a female native Hindi speaker in randomized order during a single setting. After every ten sentences, the recording was paused to ensure that the tempo and volume was not inconsistent. All the sentences were then checked by another Hindi native speaker to ensure that there was no variation in the volume and tempo for each item. For each sentence, the offset up to the verb for each item and condition was noted. The sentence length for each item prior to the verb did not vary significantly across conditions. The mean length (in ms) of the sentences prior to the critical region of the verb (i.e. sentence prefixes) were 6380 ms for the light condition, 6293 ms for the nonlight condition and 6391 ms for the anomalous condition.

In order to investigate if the prefix (i.e., the region before the light verb) was significantly different across the 3 conditions, we fit a linear regression model with the conditions as the independent variable and the length of the prefix (in ms) as the dependent variable. Treatment contrast coding was used with the non-light condition acting as the reference level. The result showed no significant difference between the baseline condition and the other conditions (Light p=0.8, t=0.24; Anomalous p=0.3, t=0.7).

Lexical Probes For each experimental sentence, we created lexical probes that were semantically unrelated to the items. We recorded individual reaction times for the probe words in a separate lexical decision task. A total of 16 native speakers of Hindi (11 males, average age=22.31) participated and carried out a lexical decision task, where a string in the Hindi Devanagari script was flashed on the screen and participants had to decide whether it was a Hindi word or a non-Hindi word. A total of 102 words (48 words and 54 non-words) were presented to the speakers in a randomized order.

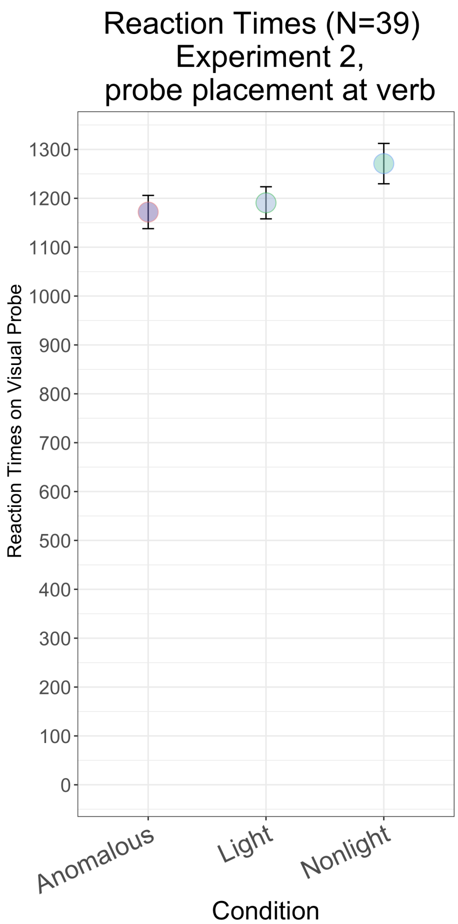
The probes were presented visually on a screen in Devanagari. Out of the 102 words we chose 15 words for the experimental sentences. In isolation, these words had a mean reaction time of 717.53 ms (SD= 33). A one-sample t-test showed that they did not differ significantly from each other (t=0, p=1). The same probe was used across all three conditions of a single item in the experiment. Each word was paired with one experimental item- the same across all conditions. As the items were presented in a Latin square, participants saw only one of the three conditions, and consequently each word probe was seen only once. We also chose 25 non-words that only appeared with the fillers. For non-words, the mean RT was 1020.87 ms (SD = 93).

2.2.3 Participants

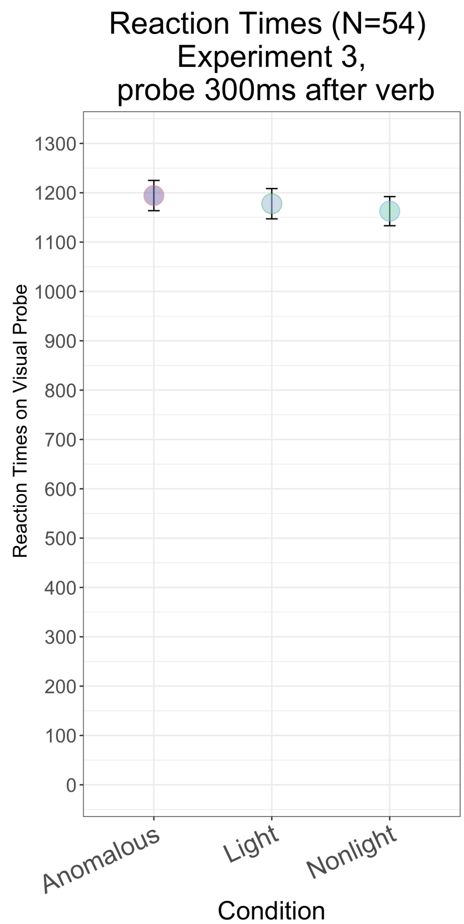
83 native speakers of Hindi participated in Experiment 2, recruited through Amazon Mechanical Turk. 39 participants (mean age=30.5) were included as part of the analysis based on performance in comprehension questions ($> 70\%$) and ($> 60\%$)accuracy at the word-non-word task.

2.2.4 Results

We used a linear mixed effects model as before, using log reaction times with condition as fixed effect and item and subject as random intercepts. Light, Non-Light and Anomalous were treatment-coded with Non-Light as the reference level. Mean reaction times for the three conditions were 1,172 ms for the anomalous, 1,190.8 ms for the light, and 1,271 ms for the non-light condition as shown in Figure 3a. In line with results from English and German, there were no significant differences between light and non-light ($t = -1.53$, $p = 0.19$) or non-light and anomalous ($t = -1.53$, $p = 0.21$).



(a) Mean Reaction times for lexical decisions in Experiment 2 (Cross modal with probe immediately after the verb



(b) Mean Reaction Times for lexical decisions in Experiment 3 (Cross-modal with probe shown 300ms after the verb.

FIGURE 3: Results for Experiments 2 and 3.

2.3 Experiment 3

This experiment used the identical paradigm as Experiment 2, i.e., the cross modal lexical decision task, and the same set of materials as before. The only difference was that the lexical probe was shown 300 ms after the verb was heard (in contrast to Experiment 2, where it was shown immediately after the verb).

2.3.1 Participants

60 Hindi native speakers (36 male, average age=20.77), who were students at IIT Delhi, participated in the experiment. Out of these, four participants were excluded due to less than 70% accuracy on the word-non-word identification task, and two due to poor performance on comprehension questions. We were left with a total of 54 Hindi speakers.

2.3.2 Results

The mean reaction times for all three conditions are shown in Figure 3b (anomalous: 1,194.3 ms, light: 1,177.8 ms, and non-light: 1,162.7 ms). Again, light, non-light, and anomalous were treatment-coded with non-light as the reference level in the analysis. The linear mixed model showed that there was no significant differences between the light and non-light condition ($t = 0.28$, $p = 0.77$), and also no difference between non-light and anomalous conditions ($t = 0.47$, $p = 0.63$).

2.4 Discussion of Experiments 2 and 3

Experiments 2 and 3 showed that when reaction times to a word probe were measured either immediately after the verb (Experiment 2), or 300ms after the verb (Experiment 3), there was no difference in reaction times to probes between the light and non-light conditions.

This lack of difference in reaction time between the light and the non-light condition fits with the experimental results of the self-paced reading study (Experiment 1), where there was no evidence for a difference in processing cost as measured by reading times between those two conditions. However, the lack of difference in reaction times to light versus non-light constructions differs from previous results using the same paradigm in English and German, where differences were found reliably with late probe placement (Piñango et al. 2006, Wittenberg and Piñango 2011), and is a useful point of cross-linguistic comparison.

Unlike in the self-paced reading task, the anomalous condition did not result in slower reaction times in either of the probe placements, although the set of items used across all three experiments was the same. The fact that the lexical decision task is not sensitive to the anomalous condition at odds with the results of Experiment 1. We have little doubt in the adequateness of our sentences, because Experiment 1 showed that the anomalous condition caused the expected slowdown. Instead, the lack of effect for anomalous sentences may be explained with reference to Wittenberg et al. (2014), which found a 'semantic P600' in response to anomalous constructions – a neural signature that has been described as reflecting a violation of overall propositional coherence, triggered by impossible, unparseable combinations (see (Kuperberg 2013, Kuperberg et al. 2020) for reviews). We would not expect such constructions that render event composition impossible to interfere with working memory later on; thus, a lack of slowdown in our working memory interference task is not completely surprising. We suggest to study this in the future.

Based on results from English and German, we trust the cross-modal decision task paradigm itself, and its sensitivity to semantically plausible but complex constructions. With this premise, the cross-linguistic differences between the light and non-light condition provide an important point of comparison for the construction in these languages (see section 3 for more discussion on this point), and we argue that taken together, the results of these three experiments imply that in Hindi, the event composition process does not result in a slowdown in reaction times at the light verb, unlike (Piñango et al. 2006, Wittenberg and Piñango 2011). This suggests that the event composition process does not seem to incur a cost for Hindi native speakers like it does for their German or English counterparts.

We also note that light verb constructions have a greater collocational frequency and greater predictability as compared to their non-light counterparts (see section 2.1.2). But we cannot be sure whether it was language-wide preference that facilitated processing, or whether the greater predictability of the individual light verb construction was more helpful, compared to previous work on German and English. In other words, it is possible that there may be context-dependent predictability effects which are reducing the computational cost of processing light verbs.

In Experiment 4 that follows, we control for the predictability of the light verb construction in order to tease apart both these frequency effects. We match the predictability between light and non-light constructions, operationalized through cloze probability. If we were to control the predictability of both light and non-light conditions, rather than keep them varied, we may be able to show more clearly the effects of event composition when measured at the verb. To avoid a floor effect in reaction times, we matched both conditions to be equally *low* in predictability.

2.5 Experiment 4

The aim of this experiment was to control both light and non-light conditions for token-based predictability, using the same paradigm as in Experiments 2 and 3, i.e. the cross-modal lexical decision task. Here, we can tease apart two factors that could have contributed to the null effects between light and non-light constructions in Experiments 1-3.

First, these results could have been due to Hindi native speakers' experience with the *language-wide systemic frequency* of the light verb construction in Hindi as a means to express predicates. If this is the case, we should find *faster* reaction times to light verbs when they are matched in cloze predictability: Hindi native speakers should be less surprised to hear a light verb construction than a non-light construction. However, if experience with the *individual construction* was driving reaction times, then we should find *slower* reaction times in the light condition, compared to non-light conditions, when cloze probabilities are matched. In this experiment, the anomalous sentences as control condition were omitted, focusing only on the comparison between the light and non-light constructions.

2.5.1 Method

This experiment also used a cross-modal lexical decision task paradigm. Participants heard sentences in either the light or non-light condition. A lexical probe was presented 300ms after the verb offset and participants had to decide whether the string was a word or a non-word. The sentences were randomized in a Latin square design and the same probe word was used in both conditions. Each participant also heard 15 (grammatical) filler sentences. The participants were asked 10 comprehension questions. Each participant heard a total of 30 sentences, half of which were experimental and the other half were fillers. In a manner similar to Experiment 2 and 3, there was a trial session that familiarized participants with the task. The presentation software used was also the same as Experiment 2 and 4 (jsPsych, version 5.0.3 (de Leeuw 2015)).

2.5.2 Materials

40 items were constructed, in the format described in Table 1, with a context sentence, a sentence containing either the light or non-light verb and a continuation. All items included the same light verb (*de* 'give'), as in Experiments 1-3.

A sentence completion task was used to calculate the cloze probabilities of the items. 16 native Hindi speakers (Mean age 33.06, female=12) were shown an incomplete sentence leading up to the light verb and were asked to complete it as naturally as possible. The items were presented using Ibex Farm (Drummond 2007). The responses were coded with respect to how often the light verb (*de* 'give') was predicted. Those items that were at chance (i.e. with a cloze probability of 50-60%) were not considered. 10 such items were removed. From the remaining 30 items, a subset of 15 low-cloze items were chosen. In the low-cloze group, both light and non-light conditions were almost equally predictable (light: 45% and non-light: 42%) with no significant difference between the two in a two-sample t-test (t=0.42, p=0.67). The remaining items were discarded.

The experimental items were recorded by a female native Hindi speaker in randomized order during a single setting. After every ten sentences, the recording was paused to ensure that the tempo and volume remained consistent. All sentences were then checked by another Hindi native speaker to ensure that there was no variation in the volume and tempo for each item. For each sentence, the offset up to the verb for each item and condition was noted. The sentence length for each item prior to the verb did not vary significantly across conditions (mean sentence length for light prior to the verb was 13333.33 ms and mean sentence length for non-light was 13226.27 ms. In order to investigate if the prefix (i.e., the region before the verb) was significantly different in the two conditions, we fit a linear regression model with the conditions as the independent variable and the length of the prefix (in ms) as the dependent variable. Treatment contrast coding was used with the non-light condition acting as the reference level. The result showed no significant difference between the baseline condition and the light condition p=0.89; t= 0.14.

2.5.3 Norming of sentences

The pre-verbal nouns in the light and non-light conditions were matched for frequency using a large corpus of 60 million tokens (Kilgarriff et al. 2010). The corpus was tokenized and tagged with parts of speech to calculate the noun frequencies and the collocational frequencies of noun and light verb.

The pre-verbal nouns were matched for frequency in the light and non-light conditions, with nouns in the light condition appearing 55.33 times per million and nouns in non-light appearing 58.88 times per million. A two-sample t-test showed no significant difference between the two conditions ($t = -0.18, p = 0.8$).

As seen in Experiment 1, the mean collocational frequency of the predicating noun and light verb was greater than the non-predicating noun and verb, which is to be expected (Mean light: 16.83 per million words, Mean non-light: 2.09 per million words). Despite the higher collocational frequency of the light condition, note that light and non-light collocation were not significantly different in terms of predictability (see Section 2.5.2).

2.5.4 Lexical Probes

The lexical probes used in this task were identical to those used in Experiments 2 and 3. A list of all probes and items can be found in the Appendix.

2.5.5 Participants

120 native Hindi speakers (Mean age 31 years) participated in the experiment on Amazon Mechanical Turk, out of which 59 remained after filtering based on performance on comprehension questions (more than 70 % correct) and accuracy in the lexical decision task performance (more than 70 % correct).

2.5.6 Results

Reaction times to probes in the light condition were on average 63 ms *faster* than probes in the the non-light condition (Mean light= 1265 ms, Mean non-light=1328 ms). This pattern is shown in Figure 4. As in the previous studies, we fit a linear mixed model predicting log reaction times from construction (light vs. non-light) as fixed effect, and item and subject as random intercepts. This model showed that the difference between reaction times to probes presented while hearing the light vs. non-light construction were significant ($t = -2.12$, $p = 0.01$).

2.5.7 Discussion of Experiment 4

The results of Experiment 4 show that when the predictability of light and non-light construction is equally low, listening to light verb constructions incurred significantly faster reaction times while making a lexical decision to unrelated probes than listening to non-light constructions.

This effect should not be due to predictability or frequency of the individual constructions. The matched cloze probability ensured that the number of possible verbs that could appear after the noun was roughly the same for both conditions. Similarly, the nouns in both conditions were matched for frequency. After controlling for these factors (particularly, cloze for this experiment), we had predicted that the effect of the co-composition process would show up in the form of longer reaction

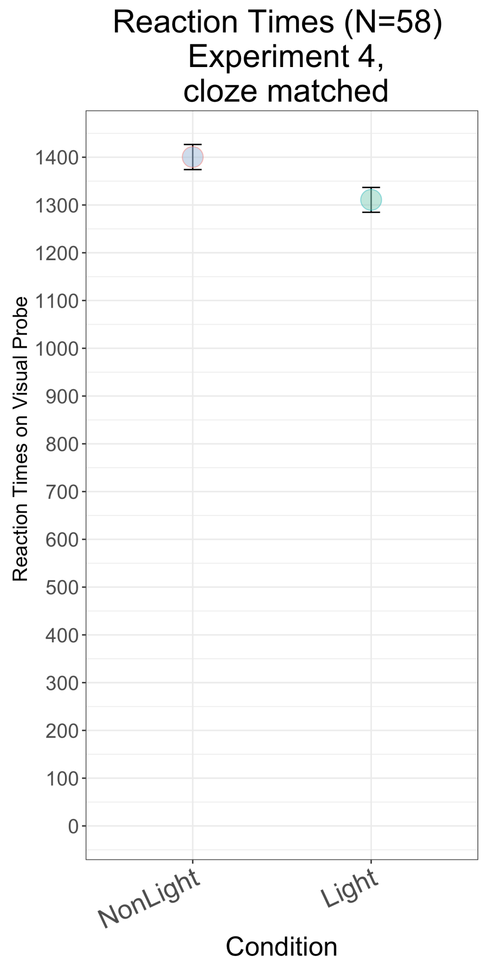


figure 4: Mean reaction times for Experiment 4 (cross-modal lexical decision task with matched cloze probabilities).

times in the light condition, while the results show the opposite pattern.

We can interpret the results in two ways: First, the collocational frequency of the predicating noun and the light verb together resulted in faster reaction times at the verb, superseding any predictability effects for the individual constructions. Another way to interpret these results is the *language-wide productivity* of the construction in Hindi: Hindi native speakers develop a greater efficiency in co-composition, where the light verb construction is being composed faster due to greater practice with this predicational strategy in the language overall.

3 General Discussion

This paper explored how the systemic, language-wide frequency of a predicational strategy affects cognitive processing, using light verb constructions in Hindi as a test case. Experiment 1 used a self-paced reading paradigm, and while people slowed down reading anomalous constructions, there was no difference in reading speed between light and non-light constructions. Experiments 2 and 3 used a cross-modal lexical decision paradigm, which has been shown to be more sensitive to semantic composition processes. However, regardless of the timing of the lexical decision task, we failed to detect any difference in reaction times to probes presented while people listened to light vs non-light constructions. Experiment 4 used the same cross-modal task, where light and non-light constructions were controlled for predictability. Crucially, in this final study, light verb constructions led to *faster* reaction times than non-light constructions. Most of these data stand in contrast to German and English data, where longer reaction times and higher processing costs were found for light verb constructions (Piñango et al. 2006, Wittenberg and Piñango 2011, Wittenberg et al. 2014, but see Wittenberg, 2013).

These results show that for speakers of Hindi, where the language-wide systemic frequency of light verb constructions is greater than in English or German, the event co-composition process does not incur a measurable computational cost when measured at the verb, or in the region immediately following the verb. We interpret these findings to imply that the process of event co-composition is facilitated by a greater exposure to the composite argument structure of light verb constructions in the language.

Thus, we interpret these datasets as evidence for language-specific effects of the systematic prevalence of a predicational strategy on cognitive processing. In Hindi, a greater proportion of predicates are expressed using complex verb phrases; light verb constructions make up more than a quarter of predicates, more than double that of English or German. Thus, Hindi native speakers have significantly more experience in processing these constructions, and this practice effect overrides any cost of co-composition.

Crucially, this interpretation hinges on two assumptions: The first of these assumptions is that the English and German data are reliable. While we have not attempted a replication of those data, we hope to do so in future work, once in-person data collection is possible again. This plan is not rooted in mistrust, but scientific prudence: In the ten years that have gone by since these data were collected, statistical methods and conventions on sample size, for instance, have changed, and so may have usage patterns, with some constructions being now perhaps more or less frequent than they used to be. Related to this point is a valid discussion on the reliability of the relatively rarely used interference paradigm, the parametrization of probe timings, and their explanatory power (see Wittenberg, 2013, for discussion). This evaluation of the paradigm is beyond the current discussion, but converging evidence from both dual tasks and Event-Related Potentials alleviates this concern (Piñango et al. 2006, Wittenberg et al. 2014, Wittenberg and Piñango 2011).

Importantly, the present results are decidedly not in line with earlier data, which brings us to the second assumption our interpretation hinges on: Namely, that in light verb construction, co-composition is indeed the explanandum. However, within the co-composition assumption, several variations on the theme are conceivable, and there are also there are several broader theoretical alternatives to consider. We discuss these in turn.

3.1 Early vs. late co-composition

The model of event co-composition that has been discussed in this paper assumes that the eventive meaning of the noun is incomplete until the integration of the light verb. This means that co-composition is not complete until after the processing of the light verb. We can refer to this as 'late' co-composition. Another possibility is a noun-driven composition account, which would predict that the process will be initiated before the verb is comprehended.

This alternative model of event composition has the noun as the sole predicator in the light verb construction (Grimshaw and Mester 1988, Kearns 1988). According to such an account, the noun is the primary predicator, while the light verb is merely a theta-marker, supplying an agent role to the subject of the clause (e.g. *Mohan* in Figure 1).

From a processing point of view, this set of accounts would predict that the co-composition process will take place early, i.e. at the noun itself, rather than after the verb is encountered. Although the noun is not usually a predicating element, in the context of the light verb construction, its eventive meaning is strongly predictive of the entire predicate. We note that this account does not rule out co-composition as a phenomenon. Rather it implies that during the real-time processing of light verb constructions the noun is so strongly predictive of the light verb that there is is a negligible cost when reaction times are measured at the verb.

Both models predict a process of co-composition, but in processing terms, the noun-driven composition account would predict the process to be initiated *earlier* than the event co-composition account, and the quality of co-composition would differ. The faster reaction times at the verb in Experiment 4 could indicate that event composition has already taken place at the noun. On the other hand, if the noun was the only predicating element, then we should have observed faster reaction times when cloze probabilities for the light condition were higher (as in Experiment 1-3). Hence our experiments do not strongly support either the late or early account, and could be compatible with both. We hope to address the time-course of co-composition in future studies.

3.2 Alternatives to event co-composition

In this paper, we have adopted event co-composition as the model to explain the process of light verb construction interpretation, but there may be alternate mechanisms contributing to the data. In this section we review some of these explanations.

Aspectual implicatures. Wittenberg and Levy (2017) have shown that the conceptualization of event duration differs between a simple transitive verb (*A kissed B*) and a light verb construction (*A gave B a kiss*). In English, a computational cost for processing light verb constructions may be attributed to incorporating these aspectual implicatures such as telicity or volitionality during the co-composition process. In Hindi, one possibility is that these aspectual implicatures are missing in light verb constructions, which could result in a reduced computational cost when measured at the verb.

One reason for this would be the absence of denominal and light verb alternations in Hindi. Only a small group of nouns in Hindi have both denominal and light verb forms, whereas in English both forms will co-exist in the language (i.e., *a kiss* vs. *to kiss*). In Hindi, the light verb construction is often the only way to express a certain meaning; no simplex verb alternative exists. Hindi native speakers must be adept at the process of using the construction to 'verbalize' new predicates into the language. Consequently, the verb in Hindi provides much less semantic content to the light verb construction as compared to German and English.

We do not have psycholinguistic evidence for the availability of aspectual implicatures in Hindi light verb constructions, but we do know, from Hindi corpus studies, that predicating nouns in light verb constructions tend to form semantically coherent groups (Sulger and Vaidya 2014). For instance, it is possible to form a light verb construction as *give a sigh/grunt/cry* but not **take a sigh/grunt/cry*: A predicating noun, such as sound emission nouns, will combine with only certain types of verbs, and not others. This suggests that there may be lexical semantic properties of the light verb that control combinatorial possibilities. If light verbs in Hindi were simply verbalizers, such restrictions should not exist- indeed they would be compatible with any noun. We note that this is still indirect evidence, and more studies need to examine the availability of these implicatures.

Collocational frequencies. It would also be possible to interpret our results as the effects of frequency and productivity alone, rather than increased practice with the co-composition process. This would amount to treating light verb constructions as any other type of multiword expression with high collocational strength, like verbal idioms or compounds, where prediction is faster due to storage as a whole construction (also see discussion in Wittenberg 2016). However, this is not supported by the results of the low cloze experiment (Experiment 4) - there is potentially more than one light verb that can appear after the noun, and yet, reaction times are faster. Additionally, on the basis of collocational strength alone, we should have also found faster reaction times in Experiment 1, which we do not. This leaves us to conclude that the process of co-composition does not appear as a processing cost as it does for English or German. Rather, this cost disappears due to language specific differences in Hindi.

Alternative syntactic and semantic configurations. As discussed above, the composite argument structure of noun and light verb is distinct from that of other verbal predicates, in Hindi and other languages. Unlike canonical one-to-one mapping between syntactic and semantic argument structure, the light verb construction needs to combine thematic roles originating from both noun and light verb. It is possible, however, that this composite argument structure is an erroneous assumption (see for discussion e.g. He and Wittenberg 2020, Wittenberg 2016); rather, light verb constructions may simply have a number of semantic roles that correspond to its syntactic arguments, where the predicating noun fills a 'metaphorical' thematic role slot. Alternatively, the predicating noun need not fill a semantic role slot at all – it simply forms a single predicate together with the light verb, and the structure has two semantic roles corresponding with two syntactic arguments. In both these scenarios, there would be no reason to believe that a composite argument structure is derived as in Figure 1, and therefore we find no computational cost associated with processing the light verb construction.

However, there is experimental evidence for co-composition in light verb constructions (again, on English). In Wittenberg and Snedeker (2014), participants were trained to categorize pictures of events according to number of thematic roles, for instance, sleeping children into a one-role category; monkeys eating bananas would be a two-role event; and a child giving an apple to a teaching would be a three-role event. In the test phase, participants also had to sort sentences containing light verb constructions containing the verb *give* (e.g., *give a kiss/kick/hug...*), base verbs (e.g., *kiss/kick/hug...*), or non-light constructions (e.g., *give a flower/plate/ticket...*). The predictions were that if the thematic argument structure of light verb constructions is constructed following surface syntactic arguments, participants will categorize light verb constructions as three-role events. If they are understood as stored constructions to describe the same as base verbs, then they would be categorized as two-role events. Results showed that sentences with light verb constructions fell between the two categories – they were categorized differently from two-role events *and* differently from three-role events, suggesting that light verb constructions may be associated with two types of argument structure simultaneously.

In a followup to this study, Wittenberg et al. (2017) conducted an eye-tracking experiment. Here, participants *implicitly* learned to classify two- and three-role sentences, without being instructed about their valency properties. Participants were able to do this successfully for non-light and base verb sentences, but when light verb constructions were encountered, they again displayed an intermediate pattern, again between the two and three-role alternatives, indicating that native speakers need to resolve two sets of thematic roles coming from the noun and light verb respectively.

None of these studies alone can answer the question of how light verb constructions are learned, stored, comprehended, and produced. However, all of them together indicate that at least in English and German, light verb constructions behave differently from canonical constructions on several different levels; and the evidence suggests that the assembly of the argument structure plays a crucial role.

3.3 Pairwise comparisons.

Experiments 2 and 3 in our study yielded no differences between the light and non-light condition. We expected to find no differences in Experiment 2 (based on the English and German results), but we found a null effect in Experiment 3 as well. In order to explain it, we examined the experiment design for the previous experiments on light verb constructions.

Interestingly, the cross-modal lexical decision task for English did not find differences between light and non-light when measured 300ms after the verb (Piñango et al. 2006). Rather, they found a pairwise difference between the light and *heavy* condition, i.e. between *Mr. Olson gave an order last night to the produce guy* (light) and *Mr. Olson typed an order last night for the produce guy* (heavy). There was no difference between the light and non-light condition at 300ms after the verb. For German, on the other hand, the light condition did result in slower reaction times than both heavy (same noun) and non-light constructions 300ms after the verb.

This seems to suggest that there could be other types of pairwise comparisons that are possible, particularly grammatically plausible ones (such as the heavy condition) rather than the semantically implausible anomalous condition used in our experiments for Hindi. Perhaps future work can examine such comparisons in more detail.

4 Conclusion

We presented four studies on comprehending Hindi light verb constructions, compared to their non-light counterparts, and anomalous sentences. In summary, there appear to be considerable differences in the speed of co-composition carried out by Hindi speakers as compared to their English and German counterparts. Our results imply that Hindi native speakers are adept at the process of understanding light verb constructions as 'verbalizing' predicates, much more so than speakers of Germanic languages. One potential explanation of these data is that the process of argument sharing is not universal, but limited to Germanic languages. However, the gist of the theoretical proposal seems to hold across languages, and was originally developed for languages like Hindi and Urdu (Butt 2010). Thus, we argue that these data provide evidence for a case of specific linguistic experiences shaping cognition: Cost disappears with practice.

Acknowledgments

We thank the audiences at SALA 2018 in Konstanz, Germany, and X-PPL 2019 in Zurich, Switzerland, for helpful comments and discussion, and gratefully acknowledge Sidharth Ranjan (IIT Delhi) for creating the custom JsPsych plugin for the presentation of the stimuli. Ashwini Vaidya was supported by the DST-CSRI (Dept of Science and Technology-Cognitive Science Research Initiative) post-doctoral fellowship at IIT Delhi (PDF-67/2015).

References

Ahmed, Tafseer, Miriam Butt, Annette Hautli, and Sebastian Sulger. 2012. A reference dependency bank for analyzing complex predicates. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12)*. `http://www.lrec-conf.org/proceedings/lrec2012/pdf/474_Paper.pdf`.

Alsina, Alex, Joan Bresnan, and Peter Sells. 1997. Complex Predicates: Structure and Theory. In *Complex Predicates*. CSLI Publications, Stanford. `https://trove.nla.gov.au/version/17082409`.

Boland, Julie E. 1997. The relationship between syntactic and semantic processes in sentence comprehension. *Language and Cognitive Processes* 12(4):423–484.

Briem, Daniela, Britta Balliel, Brigitte Rockstroh, Miriam Butt, Sabine Schulte im Walde, and Ramin Assadollahi. 2009. Distinct processing of function verb categories in the human brain. *Brain Research* 1249:173–180. `https://doi.org/10.1016/j.brainres.2008.10.027`.

Butt, Miriam. 2010. The Light Verb Jungle: Still Hacking Away. In M. Amberber, M. Harvey, and B. Baker, eds., *Complex Predicates in Cross-Linguistic Perspective*, pages 48–78. Cambridge University Press. `https://doi.org/10.1017/CBO9780511712234.004`.

Butt, Miriam, Tracy Holloway King, and Gillian Ramchand. 2008. Complex Predication:How Did the Child Pinch the Elephant? In L. Uyechi and L. Wee, eds., *Reality Exploration and Discovery: Pattern Interaction in Language and Life*. CSLI Publications, Stanford. `http://ling.uni-konstanz.de/pages/home/butt/main/papers/mo-final.pdf`.

Cuetos, Fernando, Don C. Mitchell, and Martin M.B. Corley. 1996. Parsing in different languages. In M. Carreiras, J. Garcia-Albea, and N. Sabastian-Galles, eds., *Language Processing in Spanish*. Erlbaum. `https://www.taylorfrancis.com/books/e/9780203773970/chapters/10.4324/9780203773970-11`.

Culicover, Peter W and Ray Jackendoff. 2005. *Simpler syntax*. Oxford University Press. `https://global.oup.com/academic/product/simpler-syntax-9780199271092?cc=us&lang=en&#`.

Davison, Alice. 2005. Phrasal predicates: How N combines with V in Hindi/Urdu. In T. Bhattacharya, ed., *Yearbook of South Asian Languages and Linguistics*, pages 83–116. Mouton de Gruyter. `http://www.uiowa.edu/~linguist/faculty/davison/phrasal.pdf`.

de Leeuw, J. R. 2015. jsPsych: A javascript library for creating behavioral experiments in a web browser. *Behavior Research Methods* 47(1):1–12. `https://doi.org/10.3758/s13428-014-0458-y`.

Drummond, Alex. 2007. Ibex farm. `https://spellout.net/ibexfarm/`. accessed August 8, 2020.

Durie, Mark. 1988. Verb serialization and "verbal-prepositions" in oceanic languages. *Oceanic linguistics* 27(1):1. `https://www.jstor.org/stable/3623147`.

Grillo, Nino and Joäo Costa. 2014. A novel argument for the universality of parsing principles. *Cognition* 133(1):156 – 187. `https://doi.org/10.1016/j.cognition.2014.05.019`.

Grimshaw, Jane and Armin Mester. 1988. Light verbs and theta-marking. *Linguistic Inquiry* 9(2):205–232. `https://www.jstor.org/stable/4178587`.

He, Angela Xiaoxue and Eva Wittenberg. 2020. The acquisition of event nominals and light verb constructions. *Language and Linguistics Compass* 14(2):e12363. `https://onlinelibrary.wiley.com/doi/abs/10.1111/lnc3.12363`.

Hook, Peter. 1974. *The Compound Verb in Hindi*. University of Michigan, Ann Arbor.

Jespersen, Otto. 1965. *A Modern English Grammar on Historical Principles, Part VI, Morphology*. George Allen and Unwin Ltd.

Kachru, Yamuna. 2006. *Hindi*. John Benjamins.

Kamienkowski, Juan E, Harold Pashler, Stanislas Dehaene, and Mariano Sigman. 2011. Effects of practice on task architecture: Combined evidence from interference experiments and random-walk models of decision making. *Cognition* 119(1):81–95. `https://doi.org/10.1016/j.cognition.2010.12.010`.

Kearns, Kate. 1988. Light verbs in English. Manuscript, MIT (revised 2002), `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.132.29`.

Kennedy, A., J. Pynte, W.S. Murray, and S.A. Paul. 2013. Frequency and predictability effects in the dundee corpus: an eye movement analysis. *Quarterly Journal of Experimental Psychology* 66(3):601–618. `https://doi.org/10.1080/17470218.2012.676054`.

Kilgarriff, Adam, Siva Reddy, Jan Pomikálek, and Avinesh PVS. 2010. A Corpus Factory for Many Languages. In *Proceedings of the Seventh Conference on International Language Resources and Evaluation (LREC'10)*. `http://www.lrec-conf.org/proceedings/lrec2010/pdf/79_Paper.pdf`.

Kuperberg, Gina R. 2013. The proactive comprehender: What event-related potentials tell us about the dynamics of reading comprehension. *Unraveling the behavioral, neurobiological, and genetic components of reading comprehension* pages 176–192.

Kuperberg, Gina R, Trevor Brothers, and Edward W Wlotko. 2020. A tale of two positivities and the n400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. *Journal of Cognitive Neuroscience* 32(1):12–35.

Masica, Colin. 1993. *The Indo Aryan Languages*. Cambridge University Press. `https://www.cambridge.org/in/academic/subjects/languages-linguistics/other-languages-and-linguistics/indo-aryan-languages?format=PB&isbn=9780521299442`.

McElree, Brian and Teresa Griffith. 1995. Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21(1):134.

Mitchell, Don, Fernando Cuetos, Martin Corley, and Marc Brysbaert. 1995. Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research* 24:469–488. `https://link.springer.com/article/10.1007/BF02143162`.

Mitchell, Don C. 2004. On-line methods in language processing: Introduction and historical review. In M. Carreiras and C. C. Jr., eds., *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond*, pages 15–32. New York, NY: Psychology Press. https://www.taylorfrancis.com/books/9780203509050.

Mohanan, Tara. 1997. Multidimensionality of representation- NV complex predicates in Hindi. In A. Alsina, J. Bresnan, and P. Sells, eds., *Complex Predicates*. CSLI Publications, Stanford.

Norcliffe, Elisabeth, Alice C. Harris, and T. Florian Jaeger. 2015. Cross-linguistic psycholinguistics and its critical role in theory development: early beginnings and recent advances. *Language, Cognition and Neuroscience* 30(9):1009–1032. https://doi.org/10.1080/23273798.2015.1080373.

Piñango, Maria Mercedes, J. Mack, and Ray Jackendoff. 2006. Semantic combinatorial processes in argument structure: Evidence from light-verbs. In *Proceedings of Berkeley Linguistics Society 32nd Annual Meeting*. http://dx.doi.org/10.3765/bls.v32i1.3468.

Reddy, Siva and Serge Sharoff. 2011. Cross Language POS Taggers (and other Tools) for Indian Languages: An Experiment with Kannada using Telugu Resources. In *Proceedings of the Fifth International Workshop On Cross Lingual Information Access*, pages 11–19. Chiang Mai, Thailand: Asian Federation of Natural Language Processing. https://www.aclweb.org/anthology/W11-3603.

Sadeghi, Ali Ashraf. 1993. On denominative verbs in Persian. In *Farsi Language and the Language of Science*, pages 236–246. Tehran: University Press.

Sag, Ivan A., Timothy Baldwin, Francis Bond, Ann Copestake, and Dan Flickinger. 2002. Multiword Expressions: A Pain in the neck for NLP. In *Proceedings of the 3rd International Conference on Intelligent Text Processing and Computational Linguistics (CICLing'02)*, pages 1–15. https://link.springer.com/chapter/10.1007/3-540-45715-1_1.

Seiss, Melanie. 2009. On the difference between auxiliaries, serial verbs and light verbs. In *Proceedings of the LFG09 Conference*. https://web.stanford.edu/group/cslipublications/cslipublications/LFG/14/papers/lfg09seiss.pdf.

Staub, Adrian. 2011. The effect of lexical predictability on distributions of eye fixation durations. *Psychonomic Bulletin and Review* 18(2):371–6. https://doi.org/10.3758/s13423-010-0046-9.

Sulger, Sebastian and Ashwini Vaidya. 2014. Towards Identifying Hindi/Urdu Noun Templates in Support of a Large-Scale LFG Grammar. In *Proceedings of the Fifth Workshop on South and Southeast Asian Natural Language Processing at COLING 2014*. https://www.aclweb.org/anthology/W14-5501.pdf.

Tu, Yuancheng and Dan Roth. 2011. Learning English Light Verb Constructions: Contextual or Statistical. In *Proceedings of the Workshop on Multiword Expressions (MWE 2011), 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (ACL-HLT 2011)*. https://www.aclweb.org/anthology/W11-0807.

Vaidya, Ashwini, Martha Palmer, and Bhuvana Narasimhan. 2013. Semantic roles for nominal predicates: Building a lexical resource. In *Proceedings of the 9th Workshop on Multi-word Expressions, NAACL-13*. https://www.aclweb.org/anthology/W13-1018.

Vincze, Veronika, István Nagy T, and Gabór Berend. 2011. Detecting noun compounds and light verb constructions: a contrastive study. In *Proceedings of the Workshop on Multiword Expressions: from Parsing and Generation to the Real World (MWE 2011)*. https://www.aclweb.org/anthology/W11-0817/.

Wittenberg, Eva. 2013. Paradigmenspezifische effekte subtiler semantischer manipulationen. *Linguistische Berichte* 2013(235):293–308. https://buske.de/paradigmenspezifische-effekte-subtiler-semantischer-manipulationen.html.

Wittenberg, Eva. 2016. *With light verb constructions from syntax to concepts*, vol. 7. Universitätsverlag Potsdam. https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/deliver/index/docId/8236/file/pcss07.pdf.

Wittenberg, Eva, Manizeh Khan, and Jesse Snedeker. 2017. Investigating thematic roles through implicit learning: Evidence from light verb constructions. *Frontiers in Psychology* 8:1089. https://www.frontiersin.org/article/10.3389/fpsyg.2017.01089.

Wittenberg, Eva and Roger Levy. 2017. If you want a quick kiss, make it count: How choice of syntactic construction affects event construal. *Journal of Memory and Language* 94:254 – 271. http://www.sciencedirect.com/science/article/pii/S0749596X16302480.

Wittenberg, Eva, Martin Paczynski, Heike Wiese, Ray Jackendoff, and Gina Kuperberg. 2014. The difference between "giving a rose" and "giving a kiss": Sustained neural activity to the light verb construction. *Journal of Memory and Language* 73:31–42. https://doi.org/10.1016/j.jml.2014.02.002.

Wittenberg, Eva and Maria Mercedes Piñango. 2011. Processing light verb constructions. *The Mental Lexicon* 6(3):393–413. http://dx.doi.org/10.1075/ml.6.3.03wit.

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. *Language, Cognition and Neuroscience* 29(5):635–641. https://doi.org/10.1080/01690965.2013.831918.

Appendix

Region		Context	Subject	Object	Nominal modifier	Noun	Verb	Post-verb-1	Post-verb-2	Sent end
Light vs. Nonlight	*t-value*	1.53	-0.10	-1.43	0.65	-0.76	0.45	-0.56	-0.92	-0.14
	p-value	0.12	0.91	0.15	0.51	0.44	0.65	0.57	0.35	0.89
Anomalous vs. Nonlight	*t-value*	0.24	-1.28	-1.24	0.32	1.0	4.35	2.73	0.48	0.75
	p-value	0.8	0.19	0.21	0.74	0.32	0.00001*	0.006	0.63	0.45

TABLE 4: Region-wise results for the entire sentence showing t-values (top row) and p-values (bottom row) in Experiment 1. The critical region is at the verb.

Region		Context	Subject	Object	Nominal modifier	Noun	Verb	Post-verb-1	Post-verb-2	Sent end
Light	*Mean RT*	714.12	837.92	752.61	641.66	753.38	724.82	591.28	580.27	593.32
	SD	904.04	874.07	539.24	574.82	802.61	842.96	509.89	360.2	718.02
Non Light	*Mean RT*	706.26	838.85	801.85	607.39	709.97	676.71	566.36	573.79	569.59
	SD	1820.22	945.37	809.12	500.8	480.26	577.13	406.14	344.25	615.56
Anomalous	*Mean RT*	745.14	785.58	790.86	616.92	783.61	813.32	637.5	587.79	584.82
	SD	2047.39	678.09	767.12	562.57	677.58	784.26	533.13	325.04	564.32

TABLE 5: Region-wise Mean Reaction times and standard deviations for each condition in Experiment 1. The critical region is at the verb.

1 Items in Experiment 1,2 and 3

1. अपने समय का प्रबंधन करना मुश्किल है, इसीलिए अध्यापक ने विद्यार्थी को भाषण/कैलंडर/*सिलसिला दिया और कुछ आसान उपाय भी बताये It is difficult to manage one's time, hence the teacher gave the student a *speech/calendar/*happening* and mentioned a few tips. PROBE: गाय ; cow
2. अपनी भूल को कबूल करते हुए, लंडन के अधिकारी ने ब्रिटिश अम्बेसेडर को अपना इस्तीफा/पासपोर्ट/*अफ़सोस दिया लेकिन वे फिर भी नाराज़ थे While accepting his mistake, the officer in London gave the British ambassador his *resignation/passport/*regret* but he was still upset. PROBE: चाँद; moon
3. सम्मलेन के बाद प्रोफेसर बासु ने अपने विद्यार्थियों को भारत भ्रमण करने का वचन/नक्शा/*आज़ाद दिया और साथ में उनको किस्से सुनाये After the conference, Professor Basu gave his students a *promise/map/*freedom* and recounted a few anecdotes. PROBE: चोर; thief
4. दिवाली के दिन मीरा ने एक गरीब को बाजार में इशारा/पोशाक/*फायदा दिया और उसको अपने बगल में बैठाया On the day of Diwali Meera gave a poor man a *sign/clothing/*advantage* and made him sit beside her. PROBE:साँप; snake
5. मुंबई के स्टूडियो में, मशहूर गायक राध्येश्याम ने जय को एक मौका/एल्बम/*अपराध दिया और जय को बड़ी ख़ुशी हुई In the Mumbai studio the famous singer Radheshyam gave Jay a *chance/album/*offence* and Jay was very happy. PROBE: धूल; dust
6. इस फिल्म में भगवान ने करीना के सपने में उसे दर्शन/उपन्यास/*तापमान दिया लेकिन उठने के बाद वह सपना समझ नहीं पायी In this film, God gave an *audience/novel/*weather* to Kareena in her dream, but on waking up she could not understand the dream. PROBE: जाल; web
7. ऊपर से ऑर्डर आने के बाद सीनियर इंजीनियर ने अपने विभाग को ब्रिज बनाने का निर्देश/डिजाइन/*प्रकाशन दिया और मीडिया के लिए प्रेस कांफ्रेंस की योजना बनाई After getting an order from his superiors, the senior engineer gave his department an *order/design/*publication* and made arrangements for a press conference. PROBE: नाक; nose
8. वायरस के फैलने के बाद इस संस्था ने गरीब लोगों को अस्पताल के बाहर जानकारी/जगह/घोषणा दी और उनको बड़ी राहत मिली After the virus began to spread, this organization gave the poor people *information/space/*announcement* outside the hospital and they were very relieved. PROBE: तंबाकू; tobacco

9. उसे प्रमोशन मिलने पर दादी ने रोहन को बधाई/बाइक/*पसन्द दी और पूरे बिल्डिंग को दावत भी दी After he got his promotion, grandma gave Rohan *compliments/motorcycle/*liking* and threw a party for the whole building. PROBE: कुर्सी; chair
10. आज अख़बार में लिखा था कि उस क्रांतिकारी गुट ने ही लोगों को उत्तर/हथियार/*व्यवहार दिया और उसके बाद वे जंगल में वापस चले गए Today the newspapers reported that the revolutionary group gave an *answer/weapon/*behaviour* to the people and then they went back into the jungle. PROBE: रेत; sand
11. एक निश्चित समय पर रिपोर्टर ने बाहर खड़े हुए सहायक को संकेत/पत्र/*चर्चा दिया और उसके बाद उसे कुछ समझाने लगा At a particular time, the reporter gave a *signal/letter/*argument* to his aide standing outside and then began to explain something to him: PROBE: जूता; shoe
12. क्रिकेट की ट्रेनिंग के लिए बाबूजी ने रोहन को स्टेडियम में खेलने के लिए प्रोत्साहन/किराया/*तारीफ़ दिया और अगले महीने उसका खेल देखने का वादा भी किया In order to train for cricket, Babuji gave Rohan *encouragement/fare/*praise* and promised to come see him play next month. PROBE: कमल; lotus
13. आज के कार्यक्रम में प्रधानमंत्री ने नए अध्यक्ष को कृषि सेवा का उदहारण/विभाग/प्रयास दिया और कृषि पर इस सरकार का बहुत ज़ोर है In today's event, the prime minister gave the new officer a(n) *example/division/*effort* of the agriculture ministry as this area is important to him. PROBE: कपड़े; clothes
14. उस गाँव के छोटेसे मंदिर में पंडितजी ने अभय को आश्वासन/प्रसाद/आकलन दिया क्योंकि अभय अपनी परीक्षा के लिए काफी परेशान था In the small temple in the village the priest gave Abhay *reassurance/offerings/*assessment* because Abhay was worreid about his exams. PROBE: सड़क; street
15. उस ज़िले में लोक कलाकारों ने इस आंदोलन को अपना समर्थन/संगीत/*संबंध दिया और आंदोलन ने दो महीने से काफी तेज़ी पकड़ ली है In that district, the folk artists have given this campaign their *support/music/*connection* and the campaign has intensified in the last two months. PROBE: दाल; lentil

2 Items in Experiment 4

1. कोचिंग क्लासेज बहुत पैसे लेते हैं, लेकिन कोचिंग वालों का यह फायदा है कि वे इस साल की परीक्षा में आनेवाले सवालों का अंदाजा/पेपर देते है, जिससे अच्छा अभ्यास हो सकता है. Coaching classes charge a lot of money, but one advantage of these classes is that they give a *hint/test* about the exam questions for this year and this results in better preparation. PROBE: चाँद ; moon
2. सुमन के परिवार वालों को किसी भी हालत में बेटा चाहिए था, और डिलीवरी के बाद अस्पताल कर्मचारियों ने बड़ी हिम्मत करके सुमन को सुरक्/बेटी षा दी और किसी को खबर नहीं की. Suman's family wanted a boy at any cost and after her delivery the hospital workers courageously gave Suman *protection/a girl* and made sure nobody found out. PROBE: साँप; snake.
3. सारंग गैलरी' में सुमन के कौशल्य को देखकर यूरोपियन आर्ट्स कौंसिल ने उसे बर्लिन आने का आमंत्रण/ऑप्शन दिया जहाँ वह अपनी कला को आगे बड़ा पायेगी. . After being impressed with Suman's work in the Sarang Gallery, the European Arts Council gave her an *invitation/option* to come to Berlin, where she could progress with her talent. PROBE: चोर ; thief
4. २१वी शताब्दी में भारत की सफलता को नजर में रखते हुए यू.न को भारतवासियों को एक नया दर्जा/स्वप्न देना है जिसका हमें अब एहसास हो रहा है. In the twenty-first century, keeping India's achievements in mind, the U.N. should give Indians a new *status/dream*, which we are only now beginning to realize. PROBE: धूल; dust.
5. पिछले कई सालों से भारत में इतिहास शोध की हालत देखते हुए, भारत सरकार की नयी योजना के द्वारा पैसों से इतिहासकारों को मान/आकर्षण दिया है और यह बात काफी अनपेक्षित है. Despite the sorry state of historical research in India over the last few years, the Indian government's new scheme of monetary rewards gave historians *recognition/inducement*, which was quite unexpected. PROBE: तंबाकू;tobacco.
6. उत्तराखंड के देहरादून शहर के पास 'वूडलैंड' स्कूल में सोनू के प्राध्यापक ने उसके माँ-बाप को छात्रवृत्ति के फॉर्म के लिए थोड़ी तकलीफ/फीस दी और उन्हें दो दिन के लिए वही रुकाकर रखा था. Near the city of Dehradun in Uttarakhand in Woodland school, Sonu's teacher gave his parents some *trouble/fees* and made them stay for a couple more days there. PROBE: जाल ; web.
7. इस इलाके के लोगों ने सभी अख़बारों में पार्क में पेड़ कटने के बारे में एक टिपण्णी लिखी और पार्क में हरियाली बचाने के लिए अधिकारीयों को जोर/वक्त दिया क्यूंकि यह अवैध रूप से हो रहा था. The people in this area wrote an article in

the papers about trees getting cut in the park and they gave the authorities a *bother/deadline* about the illegal nature of this work. PROBE: कुर्सी ; chair

8. परीक्षा के दिन शहर में तेज़ बारिश हुई, और देरी से आने वाले छात्रों के लिए कॉलेज ने विशेष रियायत देकर, उनको राहत/कक्षा दी जिससे वे परीक्षा पूरी कर पाये. On the day of exams, it rained heavily in the city and for the students who arrived late, the college made arrangements and gave them *assistance/(a) classroom* where they could complete their exams. PROBE: गाय; cow.
9. मेडिकल शोधकर्ताओं ने अपने शोध के जरिये एड्स जैसी गंभीर बीमारी के खिलाफ लड़ने के लिए एक नया सुझाव/हथियार दिया है जिससे मरीजों को बहुत फायदा हो सकता है. The discoveries made in medical science have given those suffering from serious diseases like AIDS a new *possibility/weapon* which will surely benefit them a great deal. PROBE: रेत; sand.
10. उस क्षेत्र के सभी जिलों में दंगा-फसाद के कारण लोग अपनी खेती बारी छोड़कर भाग गए लेकिन सरकार ने उन लोगों को न कोई सुरक्षा/भूमि दी न उन्हें मुआवज़ा मिला. In all the districts in this area people have left their homes and fields due to the riots but the government has neither given any *protection/land* to these people nor have they received any other compensation. PROBE: कमल;lotus.
11. मनु चाचा को स्कूल में पढ़ाना पसंद था, फिर भी उन्होंने नया व्यवसाय शुरू करने के लिए ज्योतिष बाबा को अपनी सहमति/कुंडली दी और अपना इस्तीफ़ा लिख दिया. Uncle Manu liked to teach in school but in order to start a new business, he gave the astrologer his *consent/horoscope* and wrote his resignation. PROBE: जूता; shoe.
12. हम घर के बाहर गाड़ी का इंतजार कर रहे थे तभी रामदासजी नजर आये और उन्होंने हमें निर्मल की शादी का न्यौता/तोहफा दिया जो एक बड़े लाल और पिले डिब्बे में रखा था. We were waiting outside the house for the car, when Ramdasji came into view and he gave us an *invitation/gift* for Nirmal's wedding which was kept in a large red and yellow box. PROBE: दाल; lentil.
13. स्पोर्ट्स मिनिस्ट्री की ऑर्डर मिलने के बाद टीम के कोच जोशी जी ने नए हॉकी टीम को आकार/कप दिया जिससे बहुत लोगों को प्रेरणा मिली है. After receiving the order from the Sports Ministry, the team coach Joshi gave the new hockey team a *shape/cup* which has given a lot of people hope. PROBE: सड़क; road.
14. दस साल के बाद इस महीने पहली बार पंचायत की भेट हुई, जिसमें ग्राम पंचायत के अलग-अलग कार्यों के लिए पंचायत प्रमुख ने एक करोड़ की पूंजी निवेश की सिफारिश/फाइल दी और आगे के काम की योजना बनाई. After 10 years, the village council met this month, where the council chief gave a *recommendation/file* for one crore rupees to take care of various activities and make a plan for the days ahead. PROBE: कपड़े; clothes.
15. कृषि विश्वविद्यालय से डिग्री प्राप्त करने के बाद ,महेशजी रामपुर गाँव गए और वहाँ उन्होंने जैविक खेती करने के लिए गाँववालों को बढावा/चैलेंज दिया और अपने चाचा की ज़मीन पर अपने आप प्रयोग करने लगे. After obtaining a degree from the agricultural university, Maheshji went to Rampur and gave the villagers *motivation/(a) challenge* to farm organically and began to experiment on his uncle's land himself. PROBE: नाक; nose.