How Long Does a Quick Kiss Take? Studying Event Duration of Light Verb Constructions Using Explicit Word Embeddings

Lin de Huybrecht, Geraint A. Wiggins
Artificial Intelligence Research Group, Vrije Universiteit Brussel
Pleinlaan 9, 1050 Brussels, Belgium
{lin.de.huybrecht, geraint.wiggins}@vub.be

Abstract

Psycholinguistic research indicates that choosing one syntactic construction over another to describe an event can influence its perceived duration: Light Verb Constructions (LVCs) such as punctive events in count syntax (*to give a kiss*) and durative events in mass syntax (*to do research*) are perceived as taking less time than their Full Verb Constructions (FVCs; *to kiss* and *to research*). Similar computational results were achieved using BERT embeddings to semantically project events onto a one-dimensional Duration scale. We reproduce and further develop this experiment with explicit word embeddings from our own co-occurrence count-based vector space. By semantically projecting 158 LVC-FVC pairs onto our Duration scale, we find that LVCs are modelled as significantly shorter than FVCs. However, we do not find an overall statistically significant difference in duration between sentences containing the target LVCs and FVCs. We demonstrate that semantic properties observed in human experiments and in BERT embeddings can also be modelled using explicit word embeddings, which have the advantage of being fully transparent. However, using transcripts from spoken conversations can be challenging when studying a specific construction: optimising the extraction of sentences containing the target expressions and composition of their meanings are to be addressed in future work.

Keywords: light verb constructions, language modelling, event construal, distributional semantics

1 Introduction

Style guides used to criticise the use of Light Verb Constructions (LVCs)[1] claiming that they are replaceable by their corresponding Full Verb Constructions (FVCs; Wittenberg, 2016, p. 5). The reasoning behind this was that both constructions express the same meaning, so one should choose the more concise form over the longer one. However, research in psycholinguistics (Wittenberg and Levy, 2017) indicates that the reasoning above is not correct: the perceived duration of an event is in fact influenced by the grammatical construction used to describe it. More specifically, according to their experiments, *punctive events*[2] in *count syntax*[3] (e.g., *to give a kiss*) and *durative events*[4] in *mass syntax*[5] (e.g., *to do research*) are perceived as taking less time than their Full Verb Construction (FVC; *to kiss* and *to research*). For durative events in count syntax (e.g., *to give a talk*), they found no significant difference in perceived duration with their transitive frame (*to talk*).

Liu and Chersoni (2023) used semantic projection (Grand et al., 2022) on contextualised embeddings from Bidirectional Encoder Representations from Transformers (BERT: Devlin et al., 2019) to test if they capture the same semantic implications in terms of duration as indicated by the human experiments of Wittenberg and Levy (2017). In line with the latter, Liu and Chersoni (2023) found that the punctive events in count syntax (e.g., *to give a kiss*; Light Verb Construction, LVC) had a significantly shorter duration than in their transitive frame (e.g., *to kiss*; Full Verb Construction, FVC). However, in contrast to Wittenberg and Levy (2017), Liu and Chersoni (2023) found no significant difference in duration when comparing durative events in mass syntax (e.g., *to give advice*; LVC) to their transitive frame (e.g., *to advise*; FVC). Furthermore, in contrast to Wittenberg and Levy

[1]Other names for Light Verb Construction (LVC) include *support verb construction*, *stretched verb construction*, *complex predicate*, *expanded predicate* and *verbo-nominal construction* (Gilquin, 2019; Ronan and Schneider, 2015).

[2]*Punctive events* are “conceptually short and bounded by a natural end point (telic) [... and] sentences in which they appear are often understood to describe several instances of the same punctive event, i.e., they are understood iteratively" (Wittenberg and Levy, 2017, p. 255).

[3]“Count syntax uses number as the uniform dimension of measurement, e.g., two cups, a dance, a jump" (Barner and Snedeker 2005, cited by Liu and Chersoni 2023, p. 367).

[4]Durative events do not have a natural endpoint (atelic) and “they do not receive an iterative reading, even if the duration of the event is explicitly extended beyond a conventionally accepted time frame" (Wittenberg and Levy, 2017, p. 255).

[5]Mass syntax is underspecified and “open to comparison using various measuring dimensions, such as volume and time, e.g., some water, some dancing, some jumping" (Barner and Snedeker 2005, cited by Liu and Chersoni, 2023, p. 367).

Proceedings of the Fifteenth Language Resources and Evaluation Conference (LREC 2026), pages 9618–9634
11–16 May 2026. ©ELRA Language Resources Association (ELRA), 2026

(2017), Liu and Chersoni (2023) found that durative events in count syntax (*to give a talk*; LVCs) were significantly shorter in duration than their transitive frame (*to talk*; FVCs).

Our work was inspired by Wittenberg and Levy (2017) and Liu and Chersoni (2023). We will refer to punctive/durative events in count/mass syntax as Light Verb Constructions (LVCs) and to transitive frames as Full Verb Constructions (FVCs; Section 2.1).

2 Motivation

Liu and Chersoni (2023) extracted sentences containing the target LVCs and FVCs from Written and Spoken BNC1994 (Consortium, 1994). They obtained contextualised BERT embeddings (Devlin et al., 2019) for these sentences using the Python MINICONS library (Misra, 2022).

Liu and Chersoni represent each FVC (e.g., *kiss*) using BERT's last layer of the verb's embedding, and each LVC (*to give a kiss*), using the contextualised embedding of the nominal (*kiss*). Although BERT embeddings do not explicitly distinguish between different parts of speech[6] (POS; e.g., the verb and noun form of *kiss*), Tenney et al. (2019) and Turton et al. (2021) demonstrated that BERT models do capture syntactic information (e.g., POS tags) and semantic features. Therefore, Liu and Chersoni's representations of the LVCs implicitly incorporate information from their surrounding context words, including the light verb itself. However, their design choice does not explicitly compose the meaning of an LVC using the light verb (*give*) together with the nominal (*kiss*).

In our work, we study whether the distinction in perceived duration of LVCs and FVCs also holds for a fully transparent co-occurrence count-based vector space (Lund and Burgess, 1996). Our vector space has a separate entry for each different part-of-speech tag (POS-tag) assigned to each lemma that occurs in the underlying corpus. We model FVCs using the `VERB` vector of the corresponding lemma and compose the meaning of each LVC by adding the vectors corresponding to the light verb and the nominal (Mitchell and Lapata, 2008; Kartsaklis and Sadrzadeh, 2013; Grefenstette and Sadrzadeh, 2011a,b; Wijnholds and Sadrzadeh, 2019).

For the underlying corpus we used Spoken BNC1994 (Consortium, 1994) together with Spoken BNC2014 (Love et al., 2017), without the written parts of the corpora. Our reasoning is that LVCs occur more frequently in spoken informal contexts. According to Wierzbicka (1982, p. 757), the "*have a V*"-construction is colloquial in all English dialects. Therefore, it is generally not used in formal contexts. She gives the example of "*have a think*" and "*have a chat*" being correct but not "**have a contemplate*" or "**have a converse*".

2.1 Light Verb Constructions

The literature on Light Verb Constructions (LVCs) does not provide us with one clear and concise answer as to what exactly constitutes an LVC. Definitions range from extremely strict to extremely permissive (Gilquin, 2019). An LVC is a construction wherein the verb does not carry most of the meaning (e.g., "Alice *did a revision* of her paper"). Its meaning is "lightend" or "semantically bleached", i.e., made less heavy than in a Full Verb Construction (FVC), e.g., "Alice *revised* her paper".

An overview of different notions of an LVC, as specified by Gilquin (2019), is depicted in Table 1. The strictest definitions only allow, for example, a combination of a fixed set of high frequency delexical verbs (*give*, *make*, *take*, *have* and *do* in the English language) with an eventative deverbal noun that is identical to a verb (e.g., *have a shower*, *give a read*...). More permissive versions also admit nouns derived from a verb (e.g., *make a choice*, *take a decision*), nouns that are semantically related to a verb (e.g., *make an effort* (try), *have a game* (play)), or even specific constructions such as *to commit a crime*. Some definitions allow the noun to be preceded by a definite article or other determiners (e.g., *some*). Others include passive constructions (Mehl, 2019).

Our definition of an LVC is more permissive than the aforementioned versions: we do not restrict the delexical verb, nor the eventative deverbal noun. We only require the event to be in count syntax (`VERB a/an NOUN`, e.g., *give a kiss*), mass syntax (`VERB NOUN`, e.g., *do research*) or definite form (`VERB the NOUN`). We do not make a distinction between definite or indefinite articles (or absence thereof), as we do not use them to construct the meaning of the LVC. We exclude constructions with determiners other than articles and constructions with passive meanings.

3 Creating our Vector Space

3.1 Motivation

We require a method for obtaining the word embeddings that allows for full transparency and control over the process. So-called explicit representations (Levy and Goldberg, 2014), such as co-

[6] In fact, BERT (Devlin et al., 2019) uses WordPiece embeddings (Wu et al., 2016), which are not guaranteed to be tokenised at the word level, but can be tokenised at subword level to account for less common words.

part of LVC	1. delexical verb		
definition	high frequency verbs	general verbs	other verbs
example(s)	give, make, take, have, do	run	to commit (a crime)

part of LVC	2. eventative deverbal noun		
definition	identical to a verb	derived from a verb	semantically related to a verb
example(s)	- have a look - give a smile - do research	- make a choice (choose) - take a decision (decide) - take action (act)	- make an effort (try) - have a game (play)

Table 1: According to Gilquin (2019), a Light Verb Construction (LVC) consists of two parts: a delexical ("semantically light") verb and an eventative deverbal ("semantically heavy") noun.

occurrence count-based techniques have been shown to exhibit similar performance as neural methods: McGregor et al. (2015) achieved results comparable to word2vec (Mikolov et al., 2013a) in a term-concept-association task; Levy and Goldberg (2014) observed the same linguistic regularities (e.g., $\mathbf{king} - \mathbf{man} + \mathbf{woman} \approx \mathbf{queen}$) often exclusively attributed to Neural Network Language Models (NNLMs), thereby arguing that the linguistic regularities found in neural embeddings are not novel, but instead well preserved as they tend to be naturally present in explicit representations.

3.2 Data Preprocessing

3.2.1 Data Cleansing

We constructed a co-occurrence count-based vector space from the Spoken British National Corpus 1994 (Spoken BNC1994; Consortium, 1994) combined with the Spoken British National Corpus 2014 (Spoken BNC2014; Love et al., 2017). Only the spoken part of Spoken BNC1994 was used (i.e., files with the `stext` XML-tag) in combination with Spoken BNC2014 (contained within a separate folder from the written corpus).

We removed all XML metadata from each file, and all punctuation, interjections, unidentified words, articles and truncated expressions, from each sentence. Only the lemmatised tokens are kept, which results in text files containing lemmatised sentences of the form "`lemma1_pos1␣lemma2_pos2␣...␣lemmaN_posN\n`", where $\texttt{pos} \in \{\texttt{ADJ}, \texttt{ADV}, \texttt{SUBST}, \texttt{VERB}, \texttt{PREP}, \texttt{PRON}\}$. The newline character at the end of the sequence of tokens indicates the end of a sentence.

Punctuation is removed as it is not important for obtaining the distributional properties of word meanings. The only distinction we make is the end of a sentence, because entropy changes significantly from one sentence to another (Reynar and Ratnaparkhi, 1997). We remove articles, truncations, interjections and unclassified words as they can skew the distributional properties.

Lastly, Spoken BNC1994 and Spoken BNC2014 are annotated with anonymised tokens for male names (`--anonnamem_SUBST`), female names (`--anonnamef_SUBST`), neuter names (`--anonnamen_SUBST`), social media (`--anonsocialmedianame_SUBST`) and places (`--anonplace_SUBST`). This anonymisation is not generalised to numbers. We decided to use the token `--anonnumeral_POS`, where $\texttt{POS} \in \{\texttt{ADJ}, \texttt{SUBST}, \texttt{VERB}, \texttt{ADV}, \texttt{PRON}\}$[7], for numbers written in digits, because for our purposes it suffices to obtain one word representation per part-of-speech tag (POS-tag) for the concept of a number. The assumption here is that numbers written as digits often indicate amounts of something.

3.3 Obtaining Frequency Counts

The lemmatised sentences are used to count the frequencies of all lemmas in the underlying corpus, i.e., `lemma_pos`, where $\texttt{pos} \in \{\texttt{ADJ}, \texttt{ADV}, \texttt{SUBST}, \texttt{VERB}, \texttt{PREP}, \texttt{PRON}\}$. The obtained frequencies are used to determine the vocabulary and construct the basis for our vector space.

3.4 Choosing a Basis and Vocabulary

To avoid obtaining a basis that contains lemmas of high-frequency non-content words, we add additional restrictions for choosing a basis on top of those presented in the data cleansing process (Section 3.2.1): we exclude prepositions, pronouns, conjugations, tokens of the form `[a-z]_POS` (e.g., `p_SUBST`) and tokens with non-alphabetic characters such as `$` and `#`. Additionally, the 50 most frequently occurring lemmas are excluded from the basis (after Wijnholds, 2020).

We include all lemmas in our vocabulary with a minimum frequency of 50 (after Wijnholds, 2020).

[7] Part-of-speech tags (POS-tags) such as `VERB`, `PRON` and `ADV` do not make much sense in the context of numerals. These anomalies are caused by errors in POS-tagging in Spoken BNC1994 and Spoken BNC2014.

This decision results in a vocabulary size of 8,588 lemmas for Spoken BNC1994 and Spoken BNC2014 combined.

3.5 Counting Co-occurrences

Using the text files resulting from the data cleansing process, we count the co-occurrences of each lemma in our vocabulary (i.e., target word of the form `lemma_POS`) with the lemmas in the basis of our vector space within a 5-token context window on both sides of the target token (after Wijnholds, 2020). We exclude sentences with less than 5 tokens.

3.6 Weighting with Positive Pointwise Mutual Information

The raw co-occurrence counts are weighted using Positive Pointwise Mutual Information (PPMI: Church and Hanks, 1990; Dagan et al., 1993), i.e., for each token T in context C, the following is computed:

$$PPMI = \max\left(\log_2\left(\frac{p(T,C)}{p(T)p(C)}\right), 0\right) \qquad (1)$$

PPMI weighting was also used by Levy and Goldberg (2014) to construct explicit vector representations, which achieved similar performance to word2vec's Skip-gram with Negative Sampling (SGNS; Mikolov et al., 2013b) on word analogy tasks of the form "x_1 is to x_2 as y_1 is to y_2", where y_2 needs to be filled in.

4 Validating the Vector Space

We validated our co-occurrence count-based vector space on multiple word similarity tasks (Table 2): RG (Rubenstein and Goodenough, 1965), MC (Miller and Charles, 1991), MEN (Bruni et al., 2014), SimLex999 (Hill et al., 2015) and WordSim353 (Finkelstein et al., 2001). We report our Spearman's Rho scores between cosine similarities of the corresponding word vectors and human similarity ratings, and compare our results to the Spearman's Rho scores of the co-occurrence count-based vector space of Wijnholds (2020). The State-of-the-Art (SOTA) scores of GloVe vectors are provided as a general context (Dobó and Csirik, 2020 cited by Wijnholds, 2020, p. 130). Additionally, we report the percentage of Out-Of-Vocabulary (OOV) word pairs of each task.

Overall, our vector space performs in the same ballpark as the co-occurrence count-based vector space of Wijnholds (2020). Our vector space outperforms the co-occurrence count-based vector space of Wijnholds (2020) in 2 out of 5 tasks, i.e., RG (0.618 compared to 0.608) and WordSim353 (0.373 compared to 0.358). Our score for the MEN task is comparable to that of Wijnholds (2020) (0.544 compared to 0.553). We obtain a much lower score for MC (0.363 compared to 0.546).

Task	OOV	ρ	WH	SOTA
RG	58.46%	0.618	0.608	0.769
WordSim353	28.61%	0.373	0.358	0.706
MC	53.57%	0.363	0.546	0.832
SimLex999	15.52%	0.126	0.259	0.406
MEN	24.37%	0.544	0.553	0.798

Table 2: Results of word similarity tasks: Our Spearman's Rho (higher is better) between cosine similarities and human similarity ratings compared to Spearman's Rho of co-occurrence count-based vector space of Wijnholds (2020, WH) and SOTA (Dobó and Csirik, 2020 cited by Wijnholds, 2020, p. 130); Out-Of-Vocabulary (OOV) word pairs.

We observed the worst performance on the SimLex999 task (0.126), which is in line with Wijnholds' (2020) score of 0.259. Furthermore, the SOTA for SimLex999 (0.406) is much lower than for the other word similarity tasks, which indicates that SimLex999 is a more difficult task than the others.

The percentage of OOV words is quite high for RG (58.46%) and MC (53.57%), which is due to the limited vocabulary size of 8,588 lemmas. Lowering the minimum word frequency required to include a lemma in the vocabulary from 50 to 3, a threshold suggested by Manning and Schütze (1999, p. 182), would almost quadruple the vocabulary size and decrease the problem of words being OOV.

5 Creating the Projection Scale

5.1 Obtaining the Duration Scale

Grand et al. (2022) proposed a semantic projection technique that uses a one-dimensional scale to represent a semantic property of a word (e.g., size, speed, loudness, etc.). This technique was used by Liu and Chersoni (2023) to project BERT embeddings (Devlin et al., 2019) representing Light Verb Constructions (LVCs) and Full Verb Constructions (FVCs) on a Duration scale ranging from short to long duration. We reconstructed that Duration scale using lemmas contained in the vocabulary of our own vector space (Table 3).

The Duration scale was constructed following (Grand et al., 2022). First, we computed the centroids representing short and long duration by computing the averages of the word vectors corresponding to the aforementioned selected words (Table 3), representing the respective ends of the scale. The

Short duration	Long duration
adjectives: brief, short, immediate, short-term **nouns:** minute, moment, second	**adjectives:** long, long-term **nouns:** age, year, decade, century

Table 3: Lemma's used to construct the Duration scale on which the vectors representing events are semantically projected (Grand et al., 2022). Based on the lemma's used in (Liu and Chersoni, 2023).

Lemma	POS	Projection	Word count
decade	noun	-17.406	**171**
century		-14.600	602
year		-6.070	$30,023$
month		-2.043	$7,780$
long	adjective	-0.860	$9,505$
week	noun	-0.547	$16,825$
day		-0.312	$24,574$
hour		1.469	$9,177$
minute		5.957	$9,145$
short	adjective	8.880	$2,491$
second	noun	11.980	$1,763$

Table 4: Semantic projection scores of duration words onto the Duration scale, sorted from longest duration to shortest duration, according to the projection scores.

line between those two centroids is the Duration scale. The semantic projections onto that line are obtained by applying the formula for scalar projection of a point onto a line (Eq. 2).

$$Proj = \frac{\overrightarrow{target} \cdot \overrightarrow{DURATION}}{||\overrightarrow{DURATION}||} \tag{2}$$

5.2 Calibrating the Duration Scale

Grand et al. (2022) demonstrated the effectiveness of their semantic projection method on GloVe embeddings (Pennington et al., 2014) and other Neural Network Language Models (NNLMs). To verify the validity of our Duration scale, we calibrate it on our co-occurrence count-based vector space by semantically projecting the words used to construct the centroids (Table 3) onto the Duration scale (Fig. 1). This projection results in higher (respectively, lower) semantic projection scores corresponding to shorter (longer) duration.

The relative order of the lemmas projected onto the Duration scale is correct (Table 4), except for the position of *decade*, which indicates a longer duration than *century*. This inconsistency can be explained by the low word count (171) of *decade* in the underlying corpus. However, removing *decade* from the Duration scale still yields correct results in the calibration task (Appendix B); only the relative position of *long* is changed from between *week* and *month* to between *month* and *year*.

6 Comparing Semantic Projections of Light and Full Verb Constructions

6.1 Light and Full Verb Constructions in Isolation

Using our Duration scale (Section 5.1), we computed the semantic projection scores of the Light Verb Constructions (LVCs) and Full Verb Constructions (FVCs) in isolation, i.e., without any surrounding context words. The semantic projection scores of the LVCs are computed by adding the vector of the light verb (e.g., `give_VERB`) and the nominal (e.g., `kiss_SUBST`; Mitchell and Lapata, 2008; Kartsaklis and Sadrzadeh, 2013; Grefenstette and Sadrzadeh, 2011a,b; Wijnholds and Sadrzadeh, 2019) and applying scalar projection Eq. (2) to the resulting vector.

We semantically project 158 LVC-FVC pairs (verbs: Table 6, projection scores: Appendix D). For each such pair this step results in a projection score for each syntactic construction. The summarising statistics indicate a tendency towards larger semantic projection values for LVCs, indicating the events in LVCs have a shorter duration than FVCs w.r.t. our Duration scale (Table 5).

We perform a paired T-test to test if there is a statistically significant difference between the semantic projection values of both constructions. First, we confirm the absence of outliers with `identify_outliers` from the R package `rstatix`. Second, we check for normality using a QQ plot (Fig. 2). Since our data meets the paired T-test assumptions, we proceed with the right-tailed paired T-test ($H_0 : \mu_{LVC-FVC} \leq 0$; $\mathrm{H_A} : \mu_{\mathrm{LVC-FVC}} > 0$). We reject the null hypothesis at significance level $\alpha = 0.05$, with a p-value smaller than 2.2×10^{-16} and $\mathrm{t}(157) = 16.58$, CI $[2.39, +\infty)$. We find a large effect size indicated by a Cohen's d of 1.40, 95% CI $[1.17, 1.64]$.

Liu and Chersoni (2023) modelled LVCs by using the contextualised BERT embedding of the nominal, without explicitly incorporating the embedding of the light verb. We demonstrate that the semantic projection method of Grand et al. (2022) also works when composing each LVC by adding the embeddings of the light verb and the nominal.

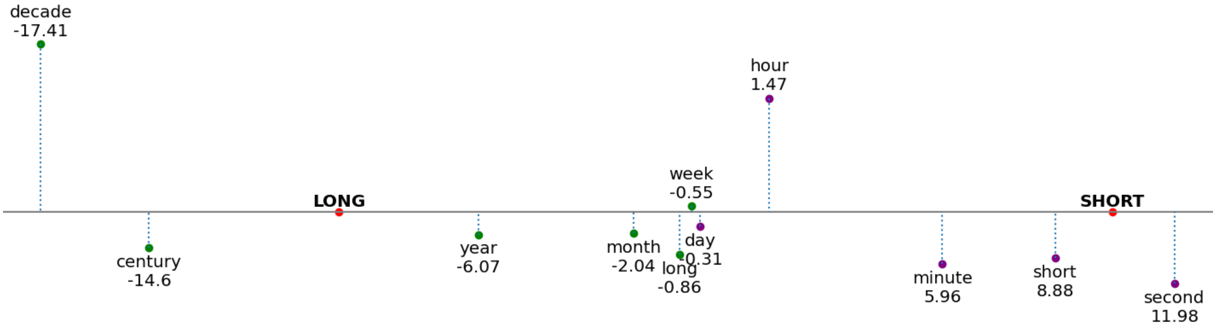

Figure 1: x-coordinates: Our semantic projection of duration words onto the Duration scale (Grand et al., 2022): **LONG** and **SHORT** are the centroids on both ends of the scale. For the y-coordinates, we used the first principal component, calculated using Principal Component Analysis. The orientation of the points makes sense for all words except for *decade*, which is explainable by its low word count in the corpus.

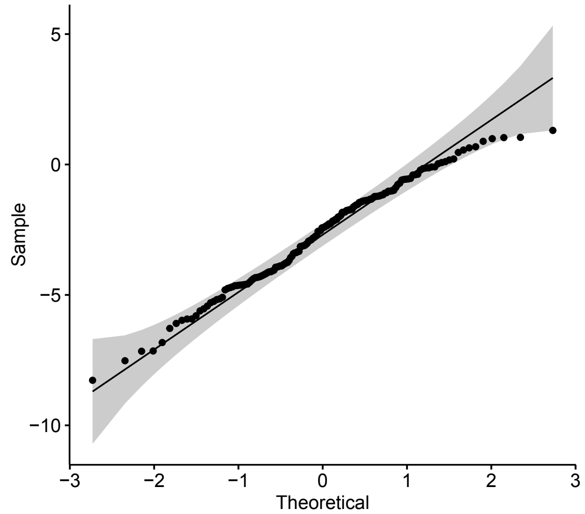

Figure 2: The semantic projection scores of Light Verb Constructions (LVCs) and Full Verb Constructions (FVCs) in isolation are normally distributed, as illustrated by the QQ plot of the differences between the scores of both constructions.

	LVC projection	FVC projection
Min.	-0.485	-2.387
Q1	2.399	0.665
Median	4.133	1.369
Mean	4.186	1.535
Std.	2.208	1.230
Q3	5.744	2.264
Max.	9.342	4.843

Table 5: Summary of semantic projection scores of plain Light Verb Constructions and Full Verb Constructions. Lower (higher) values indicate longer (shorter) duration, meaning that LVCs tend to be shorter in duration than FVCs.

6.2 Light and Full Verb Constructions in Context

We extracted all sentences from the Spoken BNC1994 and Spoken BNC2014 that contain our target Light Verb Constructions (LVCs) and Full Verb Constructions (FVCs), using the dependency parser of `Stanza` (Qi et al., 2020), a Python library which was also used by Liu and Chersoni (2023). We identified LVCs by looking for sentences that contain a target light verb (e.g., *give*) of which the object is a noun that belongs to an LVC (e.g., *kiss*). This process results in a `Text` file containing all sentences in which the target construction occurs.

Before a sentence can be projected onto the Duration scale, we need to obtain a representation for that sentence. We compose each sentence vector by adding all vectors of the corresponding tokens in each sentence obtained after data cleansing, e.g., excluding articles (Section 3.2.1). This approach is in contrast with Liu and Chersoni (2023), who represent each LVC by the contextualised BERT embedding of its nominal (e.g., "*kiss*" instead of "*give a kiss*"), thereby not explicitly composing a meaning representation for the LVC.

After applying the semantic projection Eq. (2) to the obtained sentence vector, we analyse the semantic projection scores using the Brunnel-Munzel test (Brunner and Munzel, 2000). The choice for this test is motivated by the fact that there is no good reason to assume the semantic projections are normally distributed (unpaired samples T-test), nor to assume that the semantic projections of the LVCs follow the same distribution as those of the FVCs (Wilcoxon rank sum test).

Following the advice of Karch (2023) based on Noguchi et al. (2021), we use the random permutation version of the Brunner-Munzel test implemented by the R-package `bmtest` (Karch, 2023). We conduct a one-sided Brunner-Munzel test with $H_0 : P(FVC < LVC) + \frac{1}{2}P(FVC = LVC) = 0.5$ and $H_A : P(FVC < LVC) + \frac{1}{2}P(FVC = LVC) >$

LV (VERB)	Nominal (SUBST) $\mapsto$ FVC (VERB)
commit	murder $\mapsto$ murder; offence $\mapsto$ offend
conduct	test $\mapsto$ test
do	damage $\mapsto$ damage; investigation $\mapsto$ investigate; work $\mapsto$ work
give	address $\mapsto$ address; advice $\mapsto$ advise; answer $\mapsto$ answer; assistance $\mapsto$ assist; assurance $\mapsto$ assure; battle $\mapsto$ battle; blessing $\mapsto$ bless; call $\mapsto$ call; cuddle $\mapsto$ cuddle; encouragement $\mapsto$ encourage; guidance $\mapsto$ guide; help $\mapsto$ help; hug $\mapsto$ hug; information $\mapsto$ inform; kick $\mapsto$ kick; kiss $\mapsto$ kiss; lecture $\mapsto$ lecture; look $\mapsto$ look; massage $\mapsto$ massage; miss $\mapsto$ miss; presentation $\mapsto$ present; pull $\mapsto$ pull; punch $\mapsto$ punch; push $\mapsto$ push; recognition $\mapsto$ recognise; ring $\mapsto$ ring; shake $\mapsto$ shake; shout $\mapsto$ shout; smile $\mapsto$ smile; speech $\mapsto$ speak; summary $\mapsto$ summarise; support $\mapsto$ support; talk $\mapsto$ talk; taste $\mapsto$ taste; thanks $\mapsto$ thank; treatment $\mapsto$ treat; try $\mapsto$ try; warning $\mapsto$ warn
have	argument $\mapsto$ argue; bite $\mapsto$ bite; chat $\mapsto$ chat; cough $\mapsto$ cough; accident $\mapsto$ crash; cry $\mapsto$ cry; cuddle $\mapsto$ cuddle; dance $\mapsto$ dance; doubt $\mapsto$ doubt; drink $\mapsto$ drink; feel $\mapsto$ feel; gossip $\mapsto$ gossip; kick $\mapsto$ kick; kiss $\mapsto$ kiss; laugh $\mapsto$ laugh; lead $\mapsto$ lead; look $\mapsto$ look; objection $\mapsto$ object; game $\mapsto$ play; rest $\mapsto$ rest; run $\mapsto$ run; sense $\mapsto$ sense; shave $\mapsto$ shave; smell $\mapsto$ smell; smoke $\mapsto$ smoke; swim $\mapsto$ swim; talk $\mapsto$ talk; throw $\mapsto$ throw; try $\mapsto$ try; understanding $\mapsto$ understand; walk $\mapsto$ walk; wash $\mapsto$ wash
hold	meeting $\mapsto$ meet
make	adjustment $\mapsto$ adjust; allowance $\mapsto$ allow; announcement $\mapsto$ announce; appearance $\mapsto$ appear; application $\mapsto$ apply; assessment $\mapsto$ assess; assumption $\mapsto$ assume; attempt $\mapsto$ attempt; bid $\mapsto$ bid; call $\mapsto$ call; change $\mapsto$ change; choice $\mapsto$ choose; comment $\mapsto$ comment; comparison $\mapsto$ compare; complaint $\mapsto$ complain; contribution $\mapsto$ contribute; copy $\mapsto$ copy; cut $\mapsto$ cut; decision $\mapsto$ decide; declaration $\mapsto$ declare; discovery $\mapsto$ discover; evaluation $\mapsto$ evaluate; film $\mapsto$ film; fuss $\mapsto$ fuss; improvement $\mapsto$ improve; investment $\mapsto$ invest; joke $\mapsto$ joke; mention $\mapsto$ mention; move $\mapsto$ move; offer $\mapsto$ offer; payment $\mapsto$ pay; progress $\mapsto$ progress; promise $\mapsto$ promise; proposal $\mapsto$ propose; recommendation $\mapsto$ recommend; reference $\mapsto$ refer; start $\mapsto$ start; statement $\mapsto$ state; submission $\mapsto$ submit; suggestion $\mapsto$ suggest; effort $\mapsto$ try; unity $\mapsto$ unite; use $\mapsto$ use
offer	answer $\mapsto$ answer
pay	attention $\mapsto$ attend; contribution $\mapsto$ contribute
perform	analysis $\mapsto$ analyse; comparison $\mapsto$ compare
place	advertisement $\mapsto$ advertise
play	part $\mapsto$ participate
pose	threat $\mapsto$ threaten
provide	finance $\mapsto$ finance
reach	agreement $\mapsto$ agree; decision $\mapsto$ decide
take	action $\mapsto$ act; approach $\mapsto$ approach; breath $\mapsto$ breathe; care $\mapsto$ care; decision $\mapsto$ decide; drive $\mapsto$ drive; exercise $\mapsto$ exercise; fall $\mapsto$ fall; kick $\mapsto$ kick; lead $\mapsto$ lead; look $\mapsto$ look; measurement $\mapsto$ measure; measure $\mapsto$ measure; notice $\mapsto$ notice; picture $\mapsto$ photograph; pitch $\mapsto$ pitch; punch $\mapsto$ punch; reading $\mapsto$ read; rest $\mapsto$ rest; chance $\mapsto$ risk; risk $\mapsto$ risk; shower $\mapsto$ shower; step $\mapsto$ step; taste $\mapsto$ taste; try $\mapsto$ try; walk $\mapsto$ walk
undertake	exercise $\mapsto$ exercise

Table 6: Overview of the Light Verb Constructions (LVCs) and Full Verb Constructions (FVCs) studied in this paper (LV + nominal $\mapsto$ FVC). LVCs consist of a light verb (LV), with part-of-speech (POS) tag `VERB` (e.g., `give_VERB`), and a nominal, with a `SUBST` POS tag (e.g., `kiss_SUBST`). FVCs are represented by just the verb, with POS tag `VERB` (e.g., `kiss_VERB`).

0.5. We reject the null hypothesis for 6 out of 132 LVC-FVC pairs[8] without removing outliers and reject the null hypothesis 8 times both when removing all outliers and when removing only extreme outliers (Table 7)[9]. Furthermore, for the majority of the tested LVC-FVC-pairs, we did not find a sta-

[8]Out of 144 LVC-FCC pairs for which at least 1 sentence pair was extracted for each construction of the pair, we were able to perform 132 Brunner-Munzel tests, due to insufficient data for 8 LVC-FVC pairs and a timeout error for 4 pairs.

[9]We identify outliers with `identify_outliers` from the R-package `rstatix`, which defines outliers as values outside of $[Q1 - 1.5 \times IQR, Q3 + 1.5 \times IQR]$ and extreme outliers as outside of $[Q1 - 3 \times IQR, Q3 + 3 \times IQR]$.

tistically significant difference in projection scores for $\alpha = 0.05$. However, since we conducted 132 tests, statistically, there will be more false positives than when conducting fewer statistical tests (Shaffer, 1995). Therefore, we correct for multiple testing using two correction methods at the extreme ends, i.e., leading to more false negatives or more false positives: Bonferroni correction (BF: Bonferroni, 1936) and Benjamini-Hochberg correction (BH: Benjamini and Hochberg, 1995), respectively. Under BF and BH correction, we reject no null hypotheses when we include all outliers and when removing (extreme) outliers.

Using semantic projection (Grand et al., 2022) and additive compositionality, we find no statistically significant difference in duration between sentences containing our target LVCs and FVCs, based on the number of null hypotheses rejected both with and without correction for multiple testing.

However, these results are not surprising due to a number of factors. First, the sentences we extracted from Spoken BNC1994 and Spoken BNC2014 are transcribed from spoken conversations. These sentences tend to consist of multiple clauses connected by conjunctions. Due to the nature of spoken conversational language, these clauses are not always related to each other. Moreover, for multiple target verbs, the maximum sentence length of extracted sentences is 331 tokens (Table 9). We compose the meaning of these sentences by summing the corresponding word vectors of each of these tokens, which means all tokens in the sentence are equally weighted. Additionally, clauses that are unrelated to the target LVC or FVC contribute equally to the obtained sentence vector, potentially obscuring the meaning of the target expressions and their context.

More complex compositional methods are required to adequately model the events in the sentences. Currently, it is not feasible to apply them due to the large-scale sentence extraction of over 172K sentences, with various lengths (Table 9) and grammatical constructions.

7 Conclusion & Future Work

We constructed a fully transparent co-occurrence count-based vector space from Spoken BNC1994 (Consortium, 1994) and Spoken BNC2014 (Love et al., 2017). The obtained word vectors were used to semantically project (Grand et al., 2022) Light Verb Constructions (LVCs) and Full Verb Constructions (FVCs) in isolation (i.e., without any surrounding context words) onto a Duration scale. Based on these projections, we find a statistically significant difference in duration between both constructions: LVCs are shorter in duration than FVCs. For full sentences containing the target LVCs and FVCs, we did not find a statistically significant difference in duration between sentences containing the respective expressions.

Our results are in line with the human experiments on perceived event duration of Wittenberg and Levy (2017), recalling that we grouped punctive/durative events in count/mass syntax under the term Light Verb Constructions (LVCs). Liu and Chersoni (2023) obtained similar results to Wittenberg and Levy (2017) by semantically projecting BERT embeddings (Devlin et al., 2019) onto the Duration scale. We wanted to test if this projection technique would yield similar results on word vectors obtained in a fully transparent way, namely by constructing a co-occurrence count-based vector space that uses Positive Pointwise Mutual Information (PPMI) to weight the raw co-occurrence counts. Our vector space contains a different word vector for each POS-tag that was assigned to a lemma, e.g., *kiss_VERB* and *kiss_SUBST*. Additionally, we compose the meaning of an LVC by adding the vector of the light verb (*give*) and the nominal (*kiss*), whereas Liu and Chersoni (2023) modelled the meaning of an LVC (e.g., *give a kiss*) by using the contextualised BERT embedding of the nominal, which implicitly takes into account its context.

Our current work indicates a shorter duration for LVCs than FVCs in isolation without making a distinction between event types (i.e., durative/punctive events in mass/count syntax). Future work includes making this distinction (e.g., by modelling articles) and studying the semantic shift between durative events in count syntax and their transitive frame observed by Wittenberg and Levy (2017).

In our current vector space, every target word is represented by a vector and sentence meanings are composed by summing the corresponding word vectors. However, research indicates that different parts of speech are better represented by other structures: Baroni and Zamparelli (2010) represent nouns using vectors and adjectives using matrices; Grefenstette and Sadrzadeh (2011b) represent transitive verbs with matrices. A next step would be to construct matrices for relational words such as adjectives and verbs.

Using more sophisticated methods to compose the meaning of sentences would benefit from more precise sentence extraction, i.e., extracting sentences from the corpus that contain the target LVCs or FVCs, without unrelated utterances. Optimisations could include using minimum and maximum sentence length thresholds, based on sentence length summary statistics.

We predict that combining this optimisation with more advanced compositional methods will lead to more insights about the role of LVCs and FVCs in sentence duration.

LVC-FVC	#LVC	#FVC	BM	p	BF p	BH p	Rel.eff. p	Rel.eff. CI
Including all outliers								
make appearance-appear	3	369	15.106	0.002	0.264	0.198	0.939	[0.792,1.096]
have try-try	13	8,359	3.422	0.005	0.660	0.198	0.733	[0.579,0.883]
have bite-bite	14	160	3.103	0.006	0.792	0.198	0.719	[0.566,0.868]
give address-address	31	255	2.706	0.006	0.792	0.198	0.637	[0.535,0.740]
have rest-rest	40	97	2.187	0.016	1.000	0.422	0.631	[0.511,0.747]
make move-move	35	4,787	1.746	0.044	1.000	0.968	0.583	[0.487,0.679]
Excluding all outliers								
give address-address	30	253	3.230	0.001	0.132	0.132	0.655	[0.556,0.754]
have kick-kick	4	481	9.160	0.003	0.396	0.132	0.767	[0.664,0.875]
make appearance-appear	3	365	14.921	0.003	0.396	0.132	0.939	[0.784,1.099]
have try-try	13	8,207	3.288	0.004	0.528	0.132	0.728	[0.577,0.884]
have bite-bite	14	156	2.929	0.007	0.924	0.185	0.712	[0.555,0.872]
have rest-rest	40	96	2.105	0.022	1.000	0.434	0.627	[0.505,0.746]
have argument-argue	68	365	2.059	0.023	1.000	0.434	0.572	[0.501,0.642]
play part-participate	40	50	1.902	0.027	1.000	0.446	0.617	[0.498,0.736]
Excluding extreme outliers								
have kick-kick	4	481	9.160	0.002	0.264	0.132	0.767	[0.657,0.875]
give address-address	30	253	3.230	0.002	0.264	0.132	0.655	[0.554,0.752]
make appearance-appear	3	365	14.921	0.003	0.396	0.132	0.939	[0.784,1.090]
have try-try	13	8,207	3.288	0.004	0.528	0.132	0.728	[0.575,0.879]
have bite-bite	14	156	2.929	0.007	0.924	0.185	0.712	[0.554,0.870]
have rest-rest	40	90	2.105	0.020	1.000	0.434	0.627	[0.506,0.747]
have argument-argue	68	365	2.059	0.023	1.000	0.434	0.572	[0.502,0.640]
play part-participate	40	50	1.902	0.028	1.000	0.462	0.617	[0.497,0.739]

Table 7: "Light Verb Construction (LVC)-Full Verb Construction (FVC)"-pairs for which we reject $H_0 : P(FVC < LVC) + \frac{1}{2}P(FVC = LVC) = 0.5$ (one-sided Brunner-Munzel test) under $\alpha = 0.05$. BF p and BH p: p-value under Bonferroni and Benjamini-Hochberg correction, respectively.

target verb	count	min.	Q1	median	mean	std.	Q3	max.
address	262	3	14.00	21	26.038	18.287	34.75	125
appear	385	2	12.00	20	25.434	19.274	32.00	139
bite	163	2	6.00	11	15.626	14.133	19.00	75
give address	31	4	8.50	14	18.258	15.410	26.00	85
have bite	16	4	6.75	9	12.000	8.108	13.00	32
have rest	41	2	7.00	12	19.317	18.327	23.00	68
have try	14	2	4.00	6	8.929	7.227	10.25	25
make appearance	3	6	8.00	10	8.667	2.309	10.00	10
make move	35	2	6.50	11	15.400	14.789	20.50	78
move	4,960	1	8.00	14	19.506	17.171	25.00	162
rest	97	2	11.00	17	21.887	17.783	29.00	104
try	8,729	1	9.00	16	21.852	19.506	27.00	321

Table 8: Sample sizes and summarising sentence length statistics of the extracted sentences that contain the respective target verbs, for which we reject H_0 of the one-sided Brunner-Munzel test under $\alpha = 0.05$ (including all outliers).

Group	Count	Min.	Q1	Median	Mean	Std.	Q3	Max.
combined	172,766	1	8	15	20.503	19.012	26	331
LVCs	7,205	2	9	16	21.423	18.917	27	331
FVCs	165,561	1	8	15	20.463	19.015	26	331

Table 9: Global summarising sentence length statistics, grouped by construction and combined.

8 Limitations

We did not subdivide our target verbs into durative/punctive events in mass/count syntax. In our current vector space it would not make sense to make this distinction, because our vocabulary does not include (in)definite articles. As a result, we model the Light Verb Constructions (LVCs) by composing the light verb and the nominal, leaving out the indefinite article *a* for punctive events in count syntax. This could influence the results.

Our vocabulary is relatively small, with only 8,588 lemmas, which leads to a large percentage of Out-Of-Vocabulary (OOV) words for word similarity validation tasks.

The sentences containing our target LVCs and FVCs still contain many words that often cause noise. This issue is caused by the transcription of utterances of spoken conversations. Therefore, the extracted sentences would benefit from more precise curation.

9 Acknowledgements

This research was supported by funding from the Flemish Government under the “Onderzoeksprogramma Artificiële Intelligentie (AI) Vlaanderen” programme.

10 Optional Supplementary Materials

The code for this paper and the Light Verb Construction (LVC)-Full Verb Construction (FVC) pairs are available on GitHub `https://github.com/computationalcreativitylabBrussels/LVC-event-duration`.

10.1 Ethical considerations and limitations

There are no ethical considerations to be made in this research.

11 Bibliographical References

David Barner and Jesse Snedeker. 2005. Quantity judgments and individuation: Evidence that mass nouns count. *Cognition*, 97(1):41–66.

Marco Baroni and Roberto Zamparelli. 2010. Nouns are vectors, adjectives are matrices: Representing adjective-noun constructions in semantic space. In *Proceedings of the 2010 conference on empirical methods in natural language processing*, pages 1183–1193.

Yoav Benjamini and Yosef Hochberg. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300.

Carlo Bonferroni. 1936. Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R istituto superiore di scienze economiche e commericiali di firenze*, 8:3–62.

E. Bruni, N. K. Tran, and M. Baroni. 2014. Multimodal Distributional Semantics. *Journal of Artificial Intelligence Research*, 49:1–47.

Edgar Brunner and Ullrich Munzel. 2000. The Nonparametric Behrens-Fisher Problem: Asymptotic Theory and a Small-Sample Approximation. *Biometrical Journal*, 42(1):17–25.

Kenneth Ward Church and Patrick Hanks. 1990. Word Association Norms, Mutual Information, and Lexicography. *Computational Linguistics*, 16(1):22–29.

Ido Dagan, Shaul Marcus, and Shaul Markovitch. 1993. Contextual word similarity and estimation from sparse data. In *Proceedings of the 31st annual meeting on Association for Computational Linguistics*, ACL ’93, pages 164–171, USA. Association for Computational Linguistics.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics. ArXiv:1810.04805 [cs].

András Dobó and János Csirik. 2020. A comprehensive study of the parameters in the creation and comparison of feature vectors in distributional semantic models. *Journal of Quantitative Linguistics*, 27(3):244–271.

Lev Finkelstein, Evgeniy Gabrilovich, Yossi Matias, Ehud Rivlin, Zach Solan, Gadi Wolfman, and Eytan Ruppin. 2001. Placing search in context: the concept revisited. In *Proceedings of the 10th international conference on World Wide Web*, pages 406–414, Hong Kong Hong Kong. ACM.

Gaëtanelle Gilquin. 2019. Light verb constructions in spoken L2 English: An exploratory cross-sectional study. *International Journal of Learner Corpus Research*, 5(2):181–206. Type: Journal Article.

Gabriel Grand, Idan Asher Blank, Francisco Pereira, and Evelina Fedorenko. 2022. Semantic projection recovers rich human knowledge of multiple object features from word embeddings. *Nature Human Behaviour*, 6(7):975–987.

Edward Grefenstette and Mehrnoosh Sadrzadeh. 2011a. Experimental support for a categorical compositional distributional model of meaning. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '11, pages 1394–1404, USA. Association for Computational Linguistics.

Edward Grefenstette and Mehrnoosh Sadrzadeh. 2011b. Experimenting with transitive verbs in a DisCoCat. In *Proceedings of the GEMS 2011 Workshop on GEometrical Models of Natural Language Semantics*, GEMS '11, pages 62–66, USA. Association for Computational Linguistics.

Felix Hill, Roi Reichart, and Anna Korhonen. 2015. SimLex-999: Evaluating Semantic Models With (Genuine) Similarity Estimation. *Computational Linguistics*, 41(4):665–695.

Julian D. Karch. 2023. bmtest: A Jamovi Module for Brunner–Munzel's Test—A Robust Alternative to Wilcoxon–Mann–Whitney's Test. *Psych*, 5(2):386–395. Number: 2.

Dimitri Kartsaklis and Mehrnoosh Sadrzadeh. 2013. Prior Disambiguation of Word Tensors for Constructing Sentence Vectors. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 1590–1601, Seattle, Washington, USA. Association for Computational Linguistics.

Omer Levy and Yoav Goldberg. 2014. Linguistic Regularities in Sparse and Explicit Word Representations. In *Proceedings of the Eighteenth Conference on Computational Natural Language Learning*, pages 171–180, Ann Arbor, Michigan. Association for Computational Linguistics.

Chenxin Liu and Emmanuele Chersoni. 2023. On Quick Kisses and How to Make Them Count: A Study on Event Construal in Light Verb Constructions with BERT. In *Proceedings of the 6th BlackboxNLP Workshop: Analyzing and Interpreting Neural Networks for NLP*, pages 367–378, Singapore. Association for Computational Linguistics.

Robbie Love, Claire Dembry, Andrew Hardie, Vaclav Brezina, and Tony McEnery. 2017. The Spoken BNC2014: Designing and building a spoken corpus of everyday conversations. *International Journal of Corpus Linguistics*, 22(3):319–344.

Kevin Lund and Curt Burgess. 1996. Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods, Instruments, & Computers*, 28(2):203–208.

Christopher D. Manning and Hinrich Schütze. 1999. *Foundations of statistical natural language processing*. MIT Press, Cambridge, Mass.

Stephen McGregor, Kat Agres, Matthew Purver, and Geraint A. Wiggins. 2015. From Distributional Semantics to Conceptual Spaces: A Novel Computational Method for Concept Creation. *Journal of Artificial General Intelligence*, 6(1):55–86.

Seth Mehl. 2019. Light verb semantics in the International Corpus of English: onomasiological variation, identity evidence and degrees of lightness. *English Language & Linguistics*, 23(1):55–80.

Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013a. Efficient Estimation of Word Representations in Vector Space. ArXiv:1301.3781 [cs].

Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013b. Distributed representations of words and phrases and their compositionality. *Advances in neural information processing systems*, 26.

George A. Miller and Walter G. Charles. 1991. Contextual correlates of semantic similarity. *Language and Cognitive Processes*, 6(1):1–28. _eprint: https://doi.org/10.1080/01690969108406936.

Kanishka Misra. 2022. minicons: Enabling flexible behavioral and representational analyses of transformer language models. *arXiv preprint arXiv:2203.13112*.

Jeff Mitchell and Mirella Lapata. 2008. Vector-based models of semantic composition. In *proceedings of ACL-08: HLT*, pages 236–244.

Kimihiro Noguchi, Frank Konietschke, Fernando Marmolejo-Ramos, and Markus Pauly. 2021. Permutation tests are robust and powerful at 0.5% and 5% significance levels. *Behavior Research Methods*, 53(6):2712–2724.

Jeffrey Pennington, Richard Socher, and Christopher Manning. 2014. Glove: Global Vectors for Word Representation. In *Proceedings of the 2014*

Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 1532–1543, Doha, Qatar. Association for Computational Linguistics.

Peng Qi, Yuhao Zhang, Yuhui Zhang, Jason Bolton, and Christopher D. Manning. 2020. Stanza: a Python natural language processing toolkit for many human languages. In *Proceedings of the 58th annual meeting of the association for computational linguistics: System demonstrations*.

Jeffrey C. Reynar and Adwait Ratnaparkhi. 1997. A Maximum Entropy Approach to Identifying Sentence Boundaries. In *Fifth Conference on Applied Natural Language Processing*, pages 16–19, Washington, DC, USA. Association for Computational Linguistics.

Patricia Ronan and Gerold Schneider. 2015. Determining light verb constructions in contemporary British and Irish English. *International Journal of Corpus Linguistics*, 20(3):326–354.

Herbert Rubenstein and John B. Goodenough. 1965. Contextual correlates of synonymy. *Communications of the ACM*, 8(10):627–633.

Juliet Popper Shaffer. 1995. Multiple hypothesis testing. *Annual review of psychology*, 46(1):561–584.

Ian Tenney, Dipanjan Das, and Ellie Pavlick. 2019. BERT Rediscovers the Classical NLP Pipeline. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 4593–4601, Florence, Italy. Association for Computational Linguistics.

Jacob Turton, Robert Elliott Smith, and David Vinson. 2021. Deriving Contextualised Semantic Features from BERT (and Other Transformer Model) Embeddings. In *Proceedings of the 6th Workshop on Representation Learning for NLP (RepL4NLP-2021)*, pages 248–262, Online. Association for Computational Linguistics.

Anna Wierzbicka. 1982. Why Can You Have a Drink When You Can't *Have an Eat? *Language*, 58(4):753–799.

Gijs Wijnholds. 2020. *A Compositional Vector Space Model of Ellipsis and Anaphora.* Thesis, Queen Mary University of London. Accepted: 2020-11-19T11:59:24Z.

Gijs Wijnholds and Mehrnoosh Sadrzadeh. 2019. Evaluating Composition Models for Verb Phrase Elliptical Sentence Embeddings. In *Proceedings of the 2019 Conference of the North {A}merican Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 261–271, Minneapolis, MN, USA. Association for Computational Linguistics (ACL).

Eva Wittenberg. 2016. *With light verb constructions from syntax to concepts*, volume 7. Universitätsverlag Potsdam.

Eva Wittenberg 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.

Yonghui Wu, Mike Schuster, Zhifeng Chen, Quoc V. Le, Mohammad Norouzi, Wolfgang Macherey, Maxim Krikun, Yuan Cao, Qin Gao, Klaus Macherey, Jeff Klingner, Apurva Shah, Melvin Johnson, Xiaobing Liu, Łukasz Kaiser, Stephan Gouws, Yoshikiyo Kato, Taku Kudo, Hideto Kazawa, Keith Stevens, George Kurian, Nishant Patil, Wei Wang, Cliff Young, Jason Smith, Jason Riesa, Alex Rudnick, Oriol Vinyals, Greg Corrado, Macduff Hughes, and Jeffrey Dean. 2016. Google's Neural Machine Translation System: Bridging the Gap between Human and Machine Translation. ArXiv:1609.08144 [cs].

12 Language Resource References

BNC Consortium. 1994. *British National Corpus*. University of Oxford. PID http://hdl.handle.net/20.500.14106/2554.

Love, R. and Dembry, C. and Hardie, A. and Brezina, V. and McEnery, T. 2017. *British National Corpus 2014*. Lancaster University & Cambridge University Press. URL: http://corpora.lancs.ac.uk/bnc2014/.

Appendix A XML Type Annotations in Spoken BNC1994 and Spoken BNC2014

Type of label/attribute	Spoken BNC1994	Spoken BNC2014
sentence	<s>	<u>
word	<w>	<w>
part-of-speech	"pos"	"class"
lemma	"hw"	"lemma"
punctuation	‘<c c5=TAG">CHARACTER</c>‘ TAG $\in$ {PUN,PUL,PUR,PUQ}	lemma="PUNC"
unclassified words	pos="UNC"	class="UNC"
truncated expressions	<trunc>	<trunc>
articles	pos="ART"	class="ART"
interjections	pos="INTERJ"	class="INTERJ"

Table A1: Overview of the different type annotations across Spoken BNC1994 and Spoken BNC2014.

Appendix B Duration Scale without Decade

Lemma	POS	Projection	Word count
century	noun	-15.249	602
year		-5.700	$30,023$
long	adjective	-1.774	$9,505$
month	noun	-1.668	$7,780$
week	noun	-0.279	$16,825$
day		-0.199	$24,574$
hour		1.332	$9,177$
minute		5.966	$9,145$
short	adjective	8.438	$2,491$
second	noun	11.664	$1,763$

Table B1: Semantic projection scores of duration words (Table 3, without *decade*) onto the Duration scale, sorted from longest duration to shortest duration, according to the projection scores.

Appendix C Sample Size Summary Statistics

	Incl. all outliers		Excl. extr. outliers		Excl. all outliers	
	#FVCs	#LVCs	#FVCs	#LVCs	#FVCs	#LVCs
min.	5	1	5	1	5	1
Q1	112.2	4.8	112	4	112	4
median	326	13	321	13	321	13
mean	$1,559.5$	42.6	$1,531$	41.8	$1,531$	41.8
Q3	$1,327.5$	40.3	$1,299$	40	$1,299$	40
max.	$18,076$	$1,539$	$17,718$	$1,500$	$17,718$	$1,500$

Table C1: Summary statistics of the number of semantically projected sentences that contain the target Full Verb Constructions (FVCs) and Light Verb Constructions (LVCs), by including all outliers, excluding extreme outliers, excluding all outliers. Outliers are identified as values outside of $[Q1 - 1.5 \times IQR, Q3 + 1.5 \times IQR]$ and extreme outliers as outside of $[Q1 - 3 \times IQR, Q3 + 3 \times IQR]$.

Appendix D Semantic Projection Scores of Light Verb Constructions (LVCs) and Full Verb Constructions (FVCs) in Isolation

Table D1: Semantic projection scores of plain Light Verb Constructions (LVCs) and Full Verb Constructions (FVCs), sorted alphabetically on LV and LV nominal. The semantic projection scores of the LVCs are computed by adding the vector of the verb and the nominal and applying scalar projection to the resulting vector.

LV POS: verb	LV nominal POS: noun	LVC projection POS: verb + noun	FVC POS: verb	FVC projection POS: verb
commit	murder	1.455101717430590000	murder	−2.387382879324430000
commit	offence	−0.485291267647467000	offend	0.194723992870522000
conduct	test	6.724964147699590000	test	2.838178305931350000
do	damage	1.385813429200450000	damage	2.023784577233020000
do	investigation	2.567766071332760000	investigate	2.194026662656610000
do	work	1.091595798106990000	work	−0.112903900212852000
give	address	7.361590231988260000	address	3.590987522532320000
give	advice	5.975636013798680000	advise	3.965125138993320000
give	answer	9.166180640164750000	answer	3.073197528439460000
give	assistance	7.474889958239280000	assist	2.324165382730210000
give	assurance	5.279036760584370000	assure	1.199626203714430000
give	battle	4.505119827490060000	battle	−0.121665813770481000
give	blessing	6.180452066111580000	bless	0.880230680412094000
give	call	8.768958383009250000	call	1.619409076726170000
give	cuddle	4.844436880580690000	cuddle	0.933664299296978000
give	encouragement	4.751239689789180000	encourage	0.310971660526408000
give	guidance	8.352142895930540000	guide	2.552464276930650000
give	help	6.819618415872380000	help	0.895091765953019000
give	hug	4.424002992575870000	hug	0.491652524039285000
give	information	7.222076715623650000	inform	2.581757909690910000
give	kick	8.214325667600000000	kick	3.408544069671330000
give	kiss	6.010381855851340000	kiss	0.090224012454748000
give	lecture	7.018062355282840000	lecture	−0.505334770546058000
give	look	6.848358826410900000	look	0.874599282559650000
give	massage	5.211011872236120000	massage	2.144303660555420000
give	miss	8.274927986909190000	miss	1.991637888170800000
give	presentation	6.199425839786660000	present	2.781018227198180000
give	pull	5.341191471491300000	pull	2.540255948699450000
give	punch	4.555495447986370000	punch	1.168296360859370000
give	push	5.041948384408140000	push	3.290287143731190000
give	recognition	5.678652967843600000	recognise	0.958891232706139000
give	ring	4.711415551599860000	ring	2.428343779300780000
give	shake	4.706022084755300000	shake	1.567523523817720000
give	shout	5.467354648984060000	shout	3.301090380516070000
give	smile	5.688644870403110000	smile	1.576509355671570000
give	speech	6.406332308794340000	speak	2.068177688089560000
give	summary	9.341769577344770000	summarise	1.068570328127740000
give	support	7.899347203322530000	support	3.845695878948170000
give	talk	6.479049263991560000	talk	1.386885699779320000
give	taste	4.727714391657550000	taste	0.540004082145933000
give	thanks	8.820984486848030000	thank	4.306644159249720000
give	treatment	6.217369059247920000	treat	0.605235172549837000
give	try	6.543420861224250000	try	1.363757237548050000
give	warning	7.292975017366760000	warn	0.133089682106060000

LV POS: verb	LV nominal POS: noun	LVC projection POS: verb + noun	FVC POS: verb	FVC projection POS: verb
have	accident	0.958215633191652000	crash	0.866575147413953000
have	argument	2.281713977869070000	argue	0.954023609768743000
have	bite	0.678336889008915000	bite	1.984875278295730000
have	chat	2.738897156906470000	chat	0.381379522209417000
have	cough	0.757539383979018000	cough	0.622292554398488000
have	cry	1.681922411504670000	cry	1.855502217645950000
have	cuddle	1.028840008446210000	cuddle	0.933664299296978000
have	dance	1.588694180218710000	dance	0.176007978430962000
have	doubt	0.788163492514024000	doubt	0.019270500843852800
have	drink	3.775538473912090000	drink	0.889378029274541000
have	feel	0.149736421341248000	feel	1.038316360474550000
have	game	2.174866073635940000	play	0.777740856158607000
have	gossip	0.903064481805744000	gossip	0.509463574735974000
have	kick	4.398728795465510000	kick	3.408544069671330000
have	kiss	2.194784983716860000	kiss	0.090224012454748000
have	laugh	0.974969784268691000	laugh	1.530947249694450000
have	lead	3.043369916488560000	lead	1.516773830017410000
have	look	3.032761954276420000	look	0.874599282559650000
have	objection	3.968195335858670000	object	2.753120209341340000
have	rest	2.100030864936610000	rest	3.091209807955700000
have	run	2.673625852708250000	run	1.297018574356160000
have	sense	1.290082680400100000	sense	−0.339516700151104000
have	shave	0.987946731414456000	shave	0.425193171332168000
have	smell	1.527835496148220000	smell	1.396827670452440000
have	smoke	2.512160592464430000	smoke	0.766329542085523000
have	swim	0.725571868273273000	swim	0.937461684310841000
have	talk	2.663452391857080000	talk	1.386885699779320000
have	throw	2.029566967473470000	throw	1.496649026204360000
have	try	2.727823989089760000	try	1.363757237548050000
have	understanding	1.065423206576830000	understand	0.512464156973876000
have	walk	2.266910653027970000	walk	1.041428049723320000
have	wash	1.361426598703220000	wash	1.204777771354200000
hold	meeting	7.049657698451800000	meet	0.224511520965420000
make	adjustment	2.385563344857820000	adjust	2.459317445202030000
make	allowance	2.096827343814090000	allow	1.075472100086410000
make	announcement	3.403644790715770000	announce	2.830188589507390000
make	appearance	2.438573803372200000	appear	3.479214231011390000
make	application	4.870362403732930000	apply	1.064550895641540000
make	assessment	5.467023790668610000	assess	4.011734110390730000
make	assumption	0.975308882243701000	assume	1.438218229743530000
make	attempt	2.040979580971810000	attempt	0.958607141411063000
make	bid	5.242018629879450000	bid	0.631185752625345000
make	call	6.308932342014860000	call	1.619409076726170000
make	change	4.057561428591900000	change	1.038127587173100000
make	choice	3.562444316628910000	choose	0.625198559188885000
make	comment	7.859794335372490000	comment	3.279037157909620000
make	comparison	2.513063628876640000	compare	0.563967924858208000
make	complaint	2.668381704491150000	complain	2.778364227892930000
make	contribution	2.781194177371830000	contribute	−0.970900399506873000
make	copy	8.167697147696460000	copy	3.572587774332480000
make	cut	3.625095274895600000	cut	1.517418538263630000
make	decision	3.630006109341890000	decide	1.863770121410930000
make	declaration	2.967435803696900000	declare	2.085049777012690000
make	discovery	1.875567464951490000	discover	0.155480168823076000

LV POS: verb	LV nominal POS: noun	LVC projection POS: verb + noun	FVC POS: verb	FVC projection POS: verb
make	effort	3.594975880522500000	try	1.363757237548050000
make	evaluation	3.755990593920450000	evaluate	2.206541166473060000
make	film	2.885905556665870000	film	1.501413191072940000
make	fuss	1.866293912454560000	fuss	−0.550496873584129000
make	improvement	1.879076389849170000	improve	0.435109660237059000
make	investment	1.255407306538960000	invest	0.540439642411916000
make	joke	4.032149471262500000	joke	0.926575465135115000
make	mention	4.055373195366280000	mention	2.234439236505350000
make	move	4.092889161031920000	move	2.868022199188610000
make	offer	5.177919144076730000	offer	2.264093051890650000
make	payment	3.188280918312190000	pay	1.360547736192480000
make	progress	3.689332325273100000	progress	1.407242063644080000
make	promise	1.759570796608100000	promise	2.790924027652630000
make	proposal	6.019005655544500000	propose	4.843201534748150000
make	recommendation	8.916768238960370000	recommend	3.668398200815130000
make	reference	6.920173160302280000	refer	4.581449747145320000
make	start	4.377825101074110000	start	1.034511517202650000
make	statement	5.425639347559260000	state	3.027209056207630000
make	submission	6.013622455370120000	submit	2.899185734736140000
make	suggestion	6.560899622543240000	suggest	4.133694669258660000
make	unity	2.253750844045790000	unite	1.242702045612150000
make	use	4.097822250763640000	use	2.263563488667570000
offer	answer	7.187400584009530000	answer	3.073197528439460000
pay	attention	5.253252636015310000	attend	1.681150355144690000
pay	contribution	2.358894846512820000	contribute	−0.970900399506873000
perform	analysis	3.044101336768850000	analyse	1.942524948149790000
perform	comparison	1.588621350673140000	compare	0.563967924858208000
place	advertisement	3.672083410888770000	advertise	1.675267183736450000
play	part	1.699105217739190000	participate	1.482663462233380000
pose	threat	0.949034185478059000	threaten	0.978820851690594000
provide	finance	4.168124031872290000	finance	−0.478896707078019000
reach	agreement	5.145161830027410000	agree	4.742770616027490000
reach	decision	3.448624376328770000	decide	1.863770121410930000
take	action	6.182218330070090000	act	1.852791995328540000
take	approach	4.405228957682690000	approach	0.118310970338649000
take	breath	3.943825225022790000	breathe	1.387138812648990000
take	care	4.807030061016130000	care	0.551763758842431000
take	chance	4.419248443914910000	risk	0.530794955329323000
take	decision	4.428118025196560000	decide	1.863770121410930000
take	drive	5.024593357785060000	drive	0.880275580488978000
take	exercise	6.611605523464270000	exercise	1.077594363906340000
take	fall	2.892270128303290000	fall	2.297288758707630000
take	kick	6.552411542460290000	kick	3.408544069671330000
take	lead	5.197052663483340000	lead	1.516773830017410000
take	look	5.186444701271210000	look	0.874599282559650000
take	measure	5.859111612154030000	measure	1.638405850331140000
take	measurement	4.352783234462780000	measure	1.638405850331140000
take	notice	6.125824673274340000	notice	1.375174638778210000
take	picture	4.441306064764150000	photograph	0.064167053665222000
take	pitch	5.760027933074520000	pitch	0.331311832460437000
take	punch	2.893581322846670000	punch	1.168296360859370000
take	reading	3.763819291607830000	read	2.430095930408590000
take	rest	4.253713611931400000	rest	3.091209807955700000
take	risk	3.283028449063960000	risk	0.530794955329323000

LV POS: verb	LV nominal POS: noun	LVC projection POS: verb + noun	FVC POS: verb	FVC projection POS: verb
take	shower	5.662316498738580000	shower	1.723442628486870000
take	step	3.310113919475670000	step	0.477774977247899000
take	taste	3.065800266517840000	taste	0.540004082145933000
take	try	4.881506736084550000	try	1.363757237548050000
take	walk	4.420593400022750000	walk	1.041428049723320000
undertake	exercise	5.697461099028020000	exercise	1.077594363906340000