

# Putting the argument back into argument structure constructions

Romain, Laurence

DOI:

[10.1515/cog-2021-0021](https://doi.org/10.1515/cog-2021-0021)

License:

Creative Commons: Attribution (CC BY)

*Document Version*

Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*

Romain, L 2022, 'Putting the argument back into argument structure constructions', *Cognitive Linguistics*, vol. 33, no. 1, pp. 35-64. <https://doi.org/10.1515/cog-2021-0021>

[Link to publication on Research at Birmingham portal](#)

## General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

## Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.



Laurence Romain\*

# Putting the argument back into argument structure constructions

<https://doi.org/10.1515/cog-2021-0021>

Received February 14, 2021; accepted May 20, 2021; published online June 24, 2021

**Abstract:** This paper shows that low-level generalisations in argument structure constructions are crucial to understanding the concept of alternation: low-level generalisations inform and constrain more schematic generalisations and thus constructional meaning. On the basis of an analysis of the causative alternation in English, and more specifically of the theme (i.e., the entity undergoing the event denoted by the verb), I show that each construction has its own schematic meaning. This analysis is conducted on a dataset composed of 11,554 instances of the intransitive non-causative construction and the transitive causative construction. The identification of lower-level generalisations feeds into the idea that language acquisition is organic and abstractions are formed only gradually (if at all) from exposure to input. So far, most of the literature on argument structure constructions has focused on the verb itself, and thus fails to capture these generalisations. I make up for this deficit through an in-depth analysis of the causative alternation.

**Keywords:** alternations; argument structure constructions; construction grammar; distributional semantics; vector space models

## 1 Introduction

In a paper exploring the competing motivations behind the organisation of argument structure constructions, John Du Bois concludes that “Grammars code best what speakers do most” (1985: 363); that is, language follows speakers’ experiences and perception of the world. Based on the assumption that knowledge of language is knowledge, there is no reason why there should be substantial discrepancies between our experience and conception of the world and the linguistic structures we use to describe these experiences and concepts. This idea is paramount within Langacker’s work. Langacker (1991: 282) argues that it is the

---

\*Corresponding author: Laurence Romain, Department of Modern Languages, University of Birmingham College of Arts and Law, Strathcona, Edgbaston, B15 2TT Birmingham, UK, E-mail: L.M.Y.Romain@bham.ac.uk

structure of our conception of events in terms of conceptual autonomy and dependence that motivates clausal organisation. Speakers can thus choose from a variety of argument structure constructions that will match their conceptualisation of an event (Croft 2012; Croft and Cruse 2004; Langacker 2008). Within this array of argument structure constructions, some are said to alternate. In this paper, it will be argued that to better understand the choice of constructions made by speakers, it is necessary to pay attention not only to the verbs that occur in these argument structure constructions and how they interact with constructional meaning, but also to a lower level of generalisation that includes (at least some of) the arguments used with these verbs in these constructions. It is the interaction of arguments, verbs and constructions that allows us to identify constructional meaning and more abstract generalisations as to the differences between two constructions. As Diessel (2019: 124) suggests, an adequate analysis of the interaction between verbs and constructions requires a probabilistic theory of argument realisation, which, most importantly, takes into account speakers' experience with verbs and argument structure constructions. In this first section, I will present arguments in favour of such an approach and introduce the subject of this study: the causative alternation in English. In Section 2, I will describe data collection, annotation and the methods used for analysis. Section 3 presents a selection of results and Section 4 discusses the findings.

## **1.1 Alternating argument structure constructions: lexical and constructional semantics**

The issue of alternating argument structures has been dealt with in different ways in the literature: verb-centred approaches focus on the verb and the various argument structures it can be used with whereas construction-based approaches choose to focus more on constructions themselves, and the semantics associated with these constructions. In this subsection I will show how construction-based approaches can also gain from including a finer-grained analysis of the lexical items that occur in argument structure constructions.

Put simply, lexical-projectionist approaches (e.g., Levin 1993; Levin and Rappaport-Hovav 2005) assume that a verb's semantics is sufficient to predict the various argument structure constructions it can be used in. Valency theory (Faulhaber 2011; Herbst 2011) also takes a somewhat verb-centric approach and argues that a verb's ability to occur in an argument structure construction is based on both item-specificity and broader generalisations. The general assumption behind lexical-projectionist approaches is that it is not necessary to posit more schematic structures at a higher level of abstraction. However, by not doing so,

these approaches fail to identify the characteristics associated with each argument structure construction and cannot account for the productivity of argument structure constructions. There are, however, more hybrid approaches to the issue, including Hampe and Schönefeld (2006) and Herbst (2018) which both make room for, or even insist on, the importance of collocations between verb and arguments slots.

Construction-based approaches tend to be more concerned with the semantics of each argument structure construction rather than verbs' ability to alternate. A somewhat extreme example of this stance is Goldberg's (2002) surface generalisation hypothesis which states that there is more to be gained by looking at various instances of the same construction with different verbs than by looking at various uses of the same verb in different constructions. While constructionist approaches give pride of place to constructional semantics, they also emphasise the role played by the verb in the elaboration of constructional meaning. As Goldberg (1995: 24) explains: "it is clearly not the case that the grammar works entirely top-down, with constructions simply imposing their meaning on unsuspecting verbs." Rather, the general assumption is that the central schematic meaning of these constructions is closely related to the meaning of the verb that prototypically occurs in said construction and less central meanings are extensions from this prototype. According to Goldberg (1995), argument structure constructions contribute meaning to the verb they occur with, and verbs profile a certain aspect of the basic event denoted by the construction. Stefanowitsch and Gries (2003), for example, have shown that the verb *give* is strongly attracted to the ditransitive construction, whose most central meaning is that of transfer and thus close to the meaning of *give*. In a similar vein, in a study on alternating argument structure constructions and the productivity of these constructions, Yi et al. (2019: 603) posit the Verb Anchor Hypothesis, which states that "semantic similarity to the anchor verb modulates speakers' choice of syntactic frame for these verbs." In this hypothesis, the anchor is the verb that occurs the most frequently with the frame such that it is considered representative of that frame.

Other researchers have argued, however, that it is not necessarily the case that constructional meaning is centred around a prototype. Croft (2003) suggests that constructional polysemy is better captured through family resemblances at a verb-class or even verb-specific level, rather than extensions from a single prototype. Even though they do identify *give* as prototypical of the ditransitive construction, Stefanowitsch and Gries (2003) also comment on the various classes of verbs that occur in the ditransitive and *to*-dative constructions. Gries and Stefanowitsch (2004) also identify differences in the semantics of each construction based on the classes of verbs that are most attracted to each member of an alternation. Perek (2014) makes a similar observation in his analysis of the conative construction

(e.g., *She sipped at her drink*) based on a modulation of Stefanowitsch and Gries's (2003) collexeme analysis which includes verb-class specific constructional meaning. In another vein, Gilquin (2013) focuses on verb senses rather than simply verb lemmas in her analysis of periphrastic causative constructions and finds that different verb senses are attracted to different constructions. Similarly, Bernolet and Coleman (2016) propose a collostructional analysis of the dative alternation in Dutch that takes into account the different senses of each verb studied and they show that these different senses show “quite different degrees of association to the constructions involved in a grammatical alternation” (Bernolet and Coleman 2016: 177). Overall, these lower levels of generalisation, at verb-class or verb-sense level, are assumed to provide a more accurate representation of speakers' knowledge. The argument presented in Croft (2003) reconciles the verb-centred and the construction-centred approaches, as it is quite plausible that speakers are aware both of a verb's possible distribution across various argument structure constructions and of generalisations at (a) more abstract level(s).

A finer-grained analysis that includes verb classes or verb senses is particularly relevant to the causative alternation as it is very frequent, it occurs with many different types of verbs and may thus have several related meanings. This is why I chose to work with verbs from different semantic groups, as described in Section 2. Another interesting feature of this alternation is that its two argument structure constructions, namely the intransitive non-causative construction (henceforth INCCx) and the transitive causative construction (henceforth TCCx) do not have the same number of arguments. In the next subsection, I will show how the choice of argument structure construction is thus closely related to the types (and number) of participants one includes in their description of an event.

## 1.2 The causative alternation: event conceptualisation and argument realisation

In this subsection I will show the importance of taking event conceptualisation and argument realisation into account when trying to account for speakers' choice of an argument structure construction. After all, as Langacker (1991: 286) points out, although participants can occur outside of events, “an event is conceptually dependant vis-à-vis its participants.” Before moving on to theoretical considerations, I will introduce in more detail the two constructions that make up the causative alternation. The intransitive non-causative construction (INCCx) can be illustrated by the examples in (1) and the transitive causative construction (TCCx) by the examples in (2):

- (1)
- a. The beer bottle shattered on the cement.
  - b. Pia's cheeks burned.
  - c. Reduce the tension, and the cable won't snap.
- (2)
- a. Bullets shatter bottles and glasses all around him.
  - b. A flush burns his cheeks.
  - c. Rocket Roscoe once delivered a ball so hard that it snapped a net cable at the U.S Open.

As is illustrated in these examples, the INCCx requires only one argument which is realised in subject position. The TCCx takes a minimum of two arguments: an agent, realised in subject position, and the entity undergoing the event, realised in object position. In both constructions, I will refer to the entity undergoing the event as the theme. In the examples presented in (1) and (2), the reader will notice that the themes *bottle(s)*, *cheeks* and *cable* can be found in both constructions but I will show in the next section that this is not always possible. I choose to call the participant that is shared by the two constructions the 'theme' following Langacker's (1991: 288) basic thematic roles presented in Figure 1. Langacker proposes that in an event where an entity undergoes a change of state, this entity (the theme) forms together with the predicate a minimal unit which he calls the thematic relationship. In a thematic relationship, the theme can be any of the first four thematic roles represented in Figure 1 where the different types of arrows denote the different types of events these entities can undergo: change of position (mover), internal change (patient), and mental process (experiencer).

The thematic relationship corresponds to the INCCx to the extent that the entity in subject position can undergo the event denoted by the verb without mention of an external cause, as illustrated in (1). It could be argued that the theme is simply a patient realised in subject position but patients are generally assumed to be prototypically passive (Dowty 1991) which is not entirely true of the theme in this construction, which could be construed as co-instigator (see for example Lemmens 1998) or facilitator. It is this specific property of the theme that is explored in this paper and that motivates the somewhat more neutral label

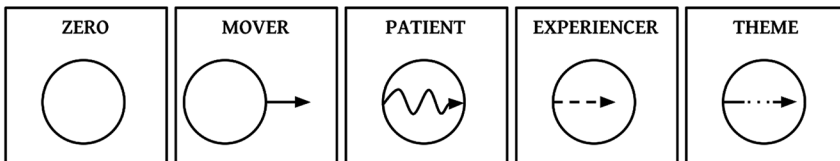
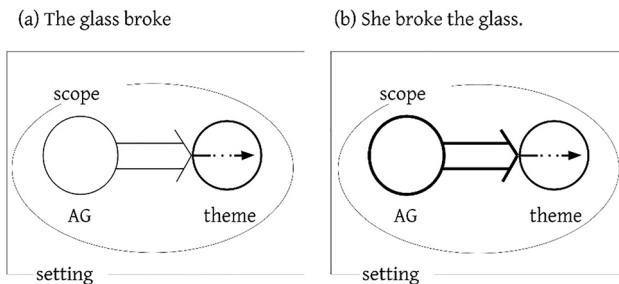


Figure 1: Basic thematic roles (Langacker 1991: 288).

‘theme’. The term theme also makes it easier to describe the entity in both the INCCx and the TCCx.

Depending on whether or not the speaker wishes to include an external cause in their description of the event, they can choose to use either the INCCx or the TCCx. These two constructions are illustrated in Figure 2, with the INCCx in (a) and the TCCx in (b). This representation shows that the glass breaking event exists pre-linguistically and, depending on how speakers conceptualise this event, they may choose one of several linguistic structures. If the speaker wants to include the cause of the event, they will use the transitive causative construction. If the speaker’s conceptualisation of the event only includes the change of state, without causation, then they will use the intransitive non-causative construction. Therefore, depending on the speaker’s conceptualisation of the event, only certain arguments will be expressed. As shown in Figure 2, the INCCx (a) only profiles the theme (in bold) whereas the TCCx (b) profiles both the theme and the agent.

The question that arises here is whether there are constraints at the level of participant/argument that may limit a speaker’s choice of construction. These constraints can be argued to be based on the semantics associated with the role at a schematic level or at a very local level, i.e., that of the interaction with the verb. The discussion around the various semantic roles that are necessary to capture generalisations in linguistic structures has been ongoing for quite a while in the literature (Boas 2003; Croft 1991, 1998, 2012; Cruse 1973; Davidse 1991, 1992; Delancey 1984; Dowty 1991; Fillmore 1977; Halliday 1967a, 1967b; Langacker 1991; Lemmens 1998, 2006; Levin and Rappaport-Hovav 2005; Talmy 2000 to name a few). Langacker argues that there are, to some extent, two extremes in the number of semantic roles that are relevant: at one end of the spectrum it can be posited that “each verb defines a distinct set of participant roles, that reflect its own unique semantic properties”; at the other end, “a role conception is arrived at by



**Figure 2:** The intransitive non-causative construction and the transitive causative construction (adapted from Langacker 1991 and Lemmens 1998).

abstracting away from the peculiarities of individual examples” (Langacker 1991: 284). In a similar vein, Croft (2012: 206–207) posits that a finer-grained analysis is preferable as it “places role designation in the semantic representation of verbs (or more precisely, verbs in particular argument structure constructions).” In other words, in order to better understand what is at stake with thematic roles, it is crucial to pay attention to their interaction with verbs in context. I will argue here that by investigating the semantic (dis)similarities between the themes that occur in the same position in one construction and the themes that occur in different positions in the two constructions, it is possible to abstract away from local generalisations to identify a more schematic meaning for each position in each construction. This view is compatible with a constructional approach in that the participant roles of a verb (lower generalisation) can be matched with the argument roles of a construction (schematic representation) (Goldberg 1995: 50).

The identification of more abstract generalisations as to the constraints on the realisation of the theme in each construction should be helpful to better understand how much information is actually shared by the two constructions and to what extent they differ.

## 2 Data collection and methods

This section will first briefly present the dataset used for this study and how it was extracted from the Corpus of Contemporary American English. Then I will describe the methods used for the analysis of the dataset: from a simple measure of Theme Overlap (2.2) to measuring the semantic (dis)similarity of themes based on distributional principles and via vector space models to place themes on a semantic map that facilitates the identification of clusters of semantically similar themes (2.3).

### 2.1 Data collection, annotation and distribution of verbs across constructions

The constructions that make up the causative alternation, INCCx and TCCx are very frequent and they occur with a large number of verbs (Levin [1993] lists 355), which makes it virtually impossible to analyse all of them. Therefore, it was necessary to restrict the dataset to certain verbs. A total of 29 verbs were selected from Levin’s list. These 29 alternating verbs are further divided into five groups, four of which were identified by Levin while the fifth, CHANGE OF TEMPERATURE was created for the purposes of this study. I chose to select groups of verbs as they can be considered as verb classes in the sense of Croft (2003) and might help identify constructional



meaning at a finer-grained level of generalisation. I extracted the data from the Corpus of Contemporary American English (COCA). The INCCx and the TCCx are virtually impossible to automatically extract with mere Part of Speech tags. Therefore, I extracted a number of sentences that contained the verbs under study and then manually annotated them for construction and theme. Table 1 presents the number of instances retained for each verb and their distribution across the two constructions. During the annotation process I had to discard a number of constructions that are similar to the INCCx and the TCCx but are not instances of these.

**Table 1:** An overview of the dataset and the distribution of its 29 verbs.

Verb group	Verb	Instances in INCCx	Instances in TCCx	Total
BREAK	<i>break</i>	121 (18.88%)	520 (81.12%)	641
	<i>crack</i>	150 (39.79%)	227 (60.21%)	377
	<i>crush</i>	2 (0.61%)	327 (99.39%)	329
	<i>shatter</i>	173 (34.67%)	326 (65.33%)	499
	<i>snap</i>	110 (33.33%)	220 (66.67%)	330
	<i>tear</i>	22 (23.91%)	70 (76.09%)	92
BEND	<i>bend</i>	167 (34.79%)	313 (65.21%)	480
	<i>crease</i>	112 (31.20%)	247(68.80%)	359
	<i>crinkle</i>	204 (61.82%)	126 (38.18%)	330
	<i>crumple</i>	185 (47.44%)	205 (52.56%)	390
	<i>fold</i>	84 (23.14%)	279 (76.86%)	363
	<i>wrinkle</i>	215 (30.67%)	486 (69.33%)	701
ROLL	<i>drop</i>	98 (27.61%)	257 (72.39%)	355
	<i>move</i>	282 (76.42%)	87 (23.58%)	369
	<i>roll</i>	86 (36.44%)	150 (63.56%)	236
	<i>slide</i>	36 (92.31%)	3 (7.69%)	39
	<i>turn</i>	81 (65.85%)	42 (34.15%)	123
GROW	<i>expand</i>	225 (30.24%)	519 (69.79%)	744
	<i>grow</i>	160 (69.78%)	69 (30.13%)	229
	<i>increase</i>	233 (30.03%)	543 (69.97%)	776
	<i>proliferate</i>	457 (93.65%)	31 (6.35%)	488
	<i>stretch</i>	114 (42.22%)	156 (57.78%)	270
	<i>thicken</i>	213 (79.18%)	56 (20.82%)	269
CHANGE OF TEMPERATURE	<i>burn</i>	281 (43.97%)	358 (56.03%)	639
	<i>chill</i>	97 (18.73%)	421 (81.27%)	518
	<i>cool</i>	149 (48.38%)	159 (51.62%)	308
	<i>freeze</i>	311 (54.95%)	255 (45.05%)	566
	<i>heat</i>	15 (4.49%)	319 (95.51%)	334
	<i>warm</i>	98 (24.50%)	302 (75.50%)	400
<b>TOTAL</b>		<b>4,481 (38.78%)</b>	<b>7,073 (61.22%)</b>	<b>11,554</b>

These constructions include the middle construction illustrated in (3) which requires an adverb to be grammatically acceptable (Goldberg 1995; Lemmens 1998, 2005; Yoshimura and Taylor 2004), objectless constructions as in (4), where the entity undergoing the event is omitted (Goldberg 2001; Lemmens 2006) and the setting construction as in (5), where the participant in subject position is not actually the agent (Davidse 1992; Guerrero-Medina 2014), which all had to be manually excluded.

- (3) This book reads well.
- (4) Crack, crush, slam and flatten more!
- (5) John broke his arm.

Table 1 shows that, at least with our 29 verbs, the TCCx is slightly more frequent than the INCCx, and that most verbs actually do alternate, although some verbs such as *slide* and *crush* barely do.

## 2.2 Theme overlap (Jaccard index)

While we see from the annotated dataset that the distribution of instances across constructions varies from one verb to the next, it is not yet clear how much overlap in meaning there is between the two constructions. I thus follow the method put forward by Lemmens (2021: 106–112) to get a clearer idea of this overlap and look at the themes that are shared by the constructions, as they are, together with the verb, the one element that is present in both constructions. To measure this overlap, we use the Jaccard index, that is, a count of the number of themes that are exclusively used with INCCx, with TCCx or that are shared by the two constructions.

The motivation behind this measure is that, even though most verbs are shared by the two constructions they might actually be used with very different themes in each construction. For example, we know that the verb *break* occurs in both constructions but some uses are restricted to one construction or the other, as shown in examples (6) and (7).

- (6) a. He broke the law.  
b. \*The law broke.
- (7) a. The day broke.  
b. \*The sun broke the day.

As is clear from (6) and (7), the theme *law* is restricted to the TCCx and the theme *day* is restricted to the INCCx when used with the verb *break*. An exaggerated assumption could be that this is true of most instances, which would subsequently

reduce the actual alternation strength of verbs. If verbs only occur in each construction with specific themes, then the amount of shared meaning between the two constructions would be drastically different from what one might expect based solely on the numbers in Table 1. A limitation, however, is that Theme Overlap only takes themes at face value. This limitation is all the more relevant that my dataset only contains 11,554 instances and it might be the case that some of the themes that appear not to be shared are actually semantically similar to themes restricted to the other construction. It is therefore necessary to identify semantic classes of themes. It will then be easier to see what kinds of themes occur in each construction, and which are shared. To achieve this, it is important to find an objective measure of semantic similarity, which is explored in the next subsection.

### 2.3 Vector space models

As Goldberg (2001: 516) observes: “Bubbles, TVs, breadsticks and hearts break in very different ways and with very different consequences.” While it is true that these things break in very different ways, they are also part of groups of entities that break in similar ways. For example, TVs break in a similar way to computers while windows and mirrors break in a similar way to each other, but very differently from TVs and computers. The fact that different things break differently leads to positing different verb senses for a verb such as *break*. The assumption here is that if we can identify clusters of semantically similar nouns, it should be possible to see whether these clusters are actually shared by the two constructions and thus whether a verb’s various senses are also used in both constructions. The collocation between theme(s) and verb(s), combined with their use in one or two specific argument structure constructions is taken as a solid ground for the elicitation of verb senses. While I could have used Behavioural Profiles (Gries and Divjak 2009) for the identification of verb senses, I chose not to for two reasons: (i) due to the size of my dataset, this would have taken an unreasonable amount of time to annotate manually and (ii) the dataset was already reduced to two specific argument structure constructions. I thus opted for a distributional semantics approach through vector space models.

Vector space models are an implementation of the principles of distributional semantics and were developed by Salton et al. (1975). The concept underlying vector space models is the representation of words (vectors) in a semantic space (for various application of vector space models see, among others, Gries and Stefanowitsch 2010; Hilpert and Perek 2015; Levshina and Heylen 2014; Padó and Lapata 2003; Peirsman et al. 2010; Perek 2016a, 2016b; Sahlgren 2008; Schütze 1992; Turney and Pantel 2010). The idea behind distributional semantics can be

Table 2: Co-occurrence matrix for *dress* and *shirt*.

	across	blue	boot	cap	cargo	core	denim	doctor	dress	example	face	front	gold	hat	wear	white
dress	0	0	0	1	0	2	1	2	0	2	1	1	2	2	2	2
shirt	2	2	2	1	2	0	1	0	2	0	1	1	2	0	3	1

roughly summed up by Firth's (1957: 11) famous observation that "you shall know a word by the company it keeps." As proposed in Lenci's (2008: 3) Distributional Hypothesis: "The degree of semantic similarity between two linguistic expressions A and B is a function of the similarity of the linguistic contexts in which A and B can appear." The expectation is thus that words that occur in similar contexts will be semantically similar. This notion of similarity is to be taken broadly and not be restricted to (near-)synonymy as such a method will consider antonyms to be semantically similar as they tend to occur in the same contexts (e.g., things that open usually also close). As such, the type of semantic similarity one obtains with vector space models is based on the words' frequently shared collocates in a given dataset, as shown in the toy co-occurrence matrix in Table 2, where the rows represent the themes and the columns their collocates. We see here that the semantically close nouns *dress* and *shirt* share certain collocates such as *denim*, *white* and the verb *wear*. Miller and Charles (1991) have shown the psychological/cognitive validity of this notion of similarity through a series of experiments.

The actual co-occurrence matrix used for the present study was taken from Perek's model (Perek 2021; see Hilpert and Perek 2015 for an illustration of this model) which contains data extracted from the Corpus of Contemporary English (COCA). It contains all the nouns that occur in COCA and the collocates were limited to the 10,000 most frequent nouns, verbs, adjectives and adverbs found in COCA within a two-word window of the target nouns; function words are deleted as they do not bear enough relevant meaning and might create too much noise. Pronouns referring to people were not retained either for this part of the analysis but they were kept in the dataset. Two additional transformations were applied: (i) the co-occurrence counts were weighed by means of Positive Pointwise Mutual Information to give more weight to collocates that co-occur more often with one theme compared to the other themes and, (ii) the number of columns was limited to 300, thereby saving only the most relevant collocates (Hilpert and Perek 2015). These transformations make the co-occurrence matrix more salient. The similarity between the word vectors is measured by calculating the distance between the different rows. This distance is measured via the 'cosine' function of the R package *lsa* (Wild 2007); the shorter the distance between two word vectors, the more similar they are. Similar themes are then clustered together via the *hclust* function in R (R-Core-Team 2013) and represented on a semantic map plotted in R. As a result, semantically similar themes will be placed close together on the map and at a distance from semantically dissimilar themes. This analysis was run for each verb in each construction, which gave two maps per verb (granted there were enough themes to compare in each use). This method presents the advantage of visualising the entire space and helps identify semantic subspaces that are more or less

populated. That is, clusters of lexemes will be more or less dense depending on the number of lexemes that populate this cluster in the semantic space.

In the next section, I will present semantic maps for a few verbs and discuss how they help visualise the distribution of each verb in the alternating constructions and what this entails for constructional meaning.

## 3 Results

This section will first present briefly the results from the Theme Overlap analysis. Then, I will present case studies for two types of verbs<sup>1</sup>: two verbs for which several senses were identified (*break* and *tear*) and three verbs whose meaning remained constant across all themes (*crease*, *crinkle* and *wrinkle*). Two dimensions are taken into account: identifying what kinds of themes are used with each or both constructions and, based on these groups of themes, identifying verb senses that may or may not be shared by the two constructions.

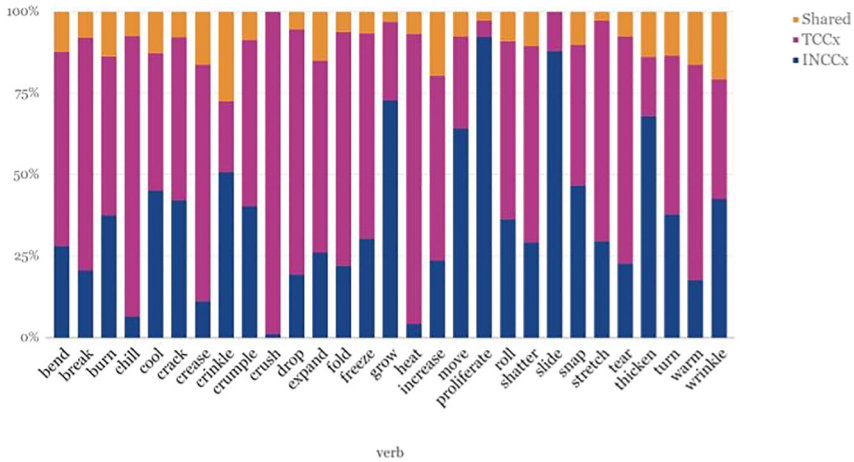
### 3.1 Theme overlap

The Theme Overlap analysis yielded results that indicate low overlap between constructions. Figure 3 shows the distribution of themes across constructions for each verb. The yellow portion represents the number of themes (types) that are actually found with both constructions in our dataset. Clearly, for many verbs, only a handful of themes are shared by the two constructions. For more than half of these verbs, the proportion of shared themes is below 10% and only three verbs have a theme overlap equal or above 20%.

This measure shows more clearly the actual distribution of verbs and their themes across the two constructions, and highlights the limited amount of overlap between the two constructions. The next sections will zoom in on the kinds of themes that are restricted or shared.

---

<sup>1</sup> I consider here that where I could identify several verb senses, the verb is polysemous and where the meaning of the verb remains constant across its uses with all themes, it is monosemous. This might be a simplistic approach to this issue, but see Evans (2009, 2018) for a more in-depth discussion of polysemy.



**Figure 3:** Theme overlap between the two constructions.

## 3.2 Vector space models

The five verbs presented here are meant to be representative of a larger tendency among verbs that alternate between the two constructions: they are found in both constructions but some of their senses are restricted to one construction or the other. Even though I cannot present all my results here, I will offer a brief overview of the 29 verbs analysed. The manual annotation of the 11,554 instances of these 29 verbs led to the identification of 4,781 different themes. Based on the clusters of themes automatically generated by the vector space model for each verb in each construction, I identified 126 different verb senses, among which only 55 are shared by the two constructions. Within these 55 shared verb senses, 12 are shared but with different themes. Finally, 31 senses are restricted to the INCCx and 40 are restricted to the TCCx. The definitions used for these verb senses come from three different sources: the Oxford English Dictionary Online (“OED Online” 2018), WordNet (Princeton University 2010) and my own interpretation when a verb sense was identified that did not match any of the definitions in these resources.

Each map presents the themes that were found with each verb in each construction. The colours were generated automatically, which is why they do not always match the clusters perfectly. As explained in Subsection 2.3, the similarity of the themes was not measured based on their use with the target verb, but in all sorts of contexts, which may lead to themes expected to undergo the event denoted by the verb in a similar way to be placed somewhat far apart on the map. The

method is not perfect, but it offers an overview of the types of themes used with each verb at a glance, which is crucial in our study.

### 3.2.1 Break

After manually annotating 1,750 instances of the verb *break*, I identified 121 instances of the INCCx and 520 instances of the TCCx, for a total of 641 instances retained. *Break* was used with 239 different themes, among which 68 were found with the INCCx and 190 with the TCCx, for a theme overlap of 7.95%. Figure 4 shows the themes used with the INCCx and Figure 5 the themes used with the TCCx. I manually added a dashed line on each map to show the separation between concrete themes from abstract themes. As is obvious from a quick glance at the two maps, there are many more abstract themes found with the TCCx than with the INCCx. In the INCCx, we find the NEWS family such as *news*, *story* and *scandal* and the PEOPLE family with *man* and *folk*. These two families of themes are actually shared by the two constructions. The abstract themes found in the TCCx can be grouped under three different families: LAW (including *curfew*, *law* and *rule*), PROMISE (as in *promise*, *oath* and *pledge*) and HABIT & SITUATION (*habit*, *routine* or *engagement*). Interestingly, these themes are actually restricted to the TCCx and cannot appear in the INCCx.

Based on these semantic maps, I isolated eight different senses of *break*, which are summed up in Table 3. In this table, the senses that are specific to the INCCx are on the right, senses restricted to the TCCx on the left and shared senses are in the middle. Among these eight senses, half are shared by the two constructions. The first sense, “(cause to) become broken”, is found with a rather large variety of themes and specifically with bones, as illustrated in (8).

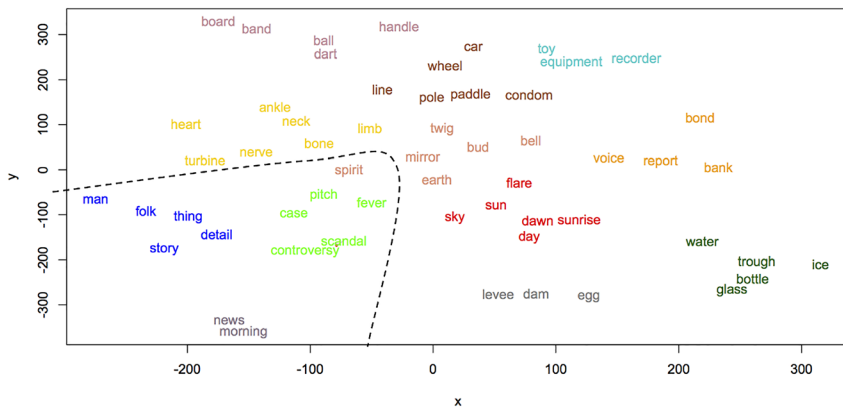


Figure 4: A semantic map of the themes that occur with *break* in the INCCx.



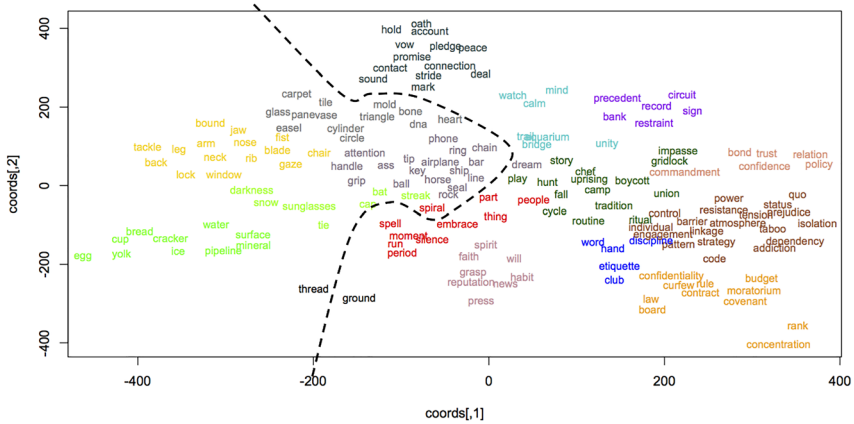


Figure 5: A semantic map of the themes that occur with *break* in the TCCx.

(8) You attacked me from behind and broke my arm this day [...]

Another interesting example within this group is *water* which is restricted to the INCCx as illustrated in (9).

(9) Juanita Massie can recall her baby’s kicks inside her belly, how her water broke [...]

Table 3: Shared and construction-specific verb senses of *break*.

<i>break</i>		
Intransitive non-causative construction	Shared	Transitive causative construction
	1.1. BONES	
	1.2. various themes	
1.3. <i>water</i>	2. MOTIVE-POWERED DEVICES	
	3. PEOPLE & MIND STATES,	
	3.1. <i>heart</i>	
	4. NEWS	
5. DAY & ATMOSPHERIC ELEMENTS		6. LAW
		7. PROMISE
		8. HABIT & SITUATION

The second verb sense identified for *break* is found with themes that refer to MOTIVE-POWERED DEVICES, and in which case the item becomes or is made inoperative, as illustrated in (10).

- (10) MOTIVE-POWERED DEVICES:
- a. My tape recorder broke, but I'm going to get it fixed tomorrow.
  - b. The intent here is for me to learn flight-testing and not break the airplane while I do [...]

The next two senses are found with more abstract themes: PEOPLE & MIND STATES and NEWS. These two senses are found with both constructions, as illustrated in (11) and (12).

- (11) PEOPLE & MIND-STATES, (cause to) lose emotional control
- a. Did police put their fists in women's faces so the husbands' spirits would break along with the women's jaws?
  - b. New York breaks a lot of people, but Spree seems to be thriving there.
- (12) NEWS
- a. After the infidelity scandal broke, paparazzi showed James' every move.
  - b. [...] our White House correspondent that broke the news this morning about the departure of Jean Bertrand-Aristide.

The last four senses are restricted to one construction. The first of these four, "appear" is found with DAY & ATMOSPHERIC ELEMENTS and is restricted to the INCCx, as shown in (13). In the Oxford English Dictionary, the definition of *break* for this sense is *To burst out of darkness, begin to shine; as of the day, morning, daylight*. Lemmens (2006) argues that there is actually a now lost motion sense with this particular sense of *break*, and that the entity in subject position is not actually the entity that undergoes the event denoted by the verb. The assumption is that themes such as *day, sun or dawn* actually do the breaking, i.e., they break the darkness. Considering the opacity of this interpretation, I chose to keep this sense as part of the INCCx.

- (13) [...] as dawn breaks in Afghanistan [...]

The last three groups are restricted to the TCCx and are all found with abstract themes. Sense 6 is found with themes such as *law* and *rule* and corresponds to *infringe*. Sense 7 is closely related and is found with themes such as *promise* and

*oath*, for these the meaning of *break* is similar to *betray*. Finally the last sense is found with themes that refer to habits and certain situations, as illustrated in (14) and (15) respectively.

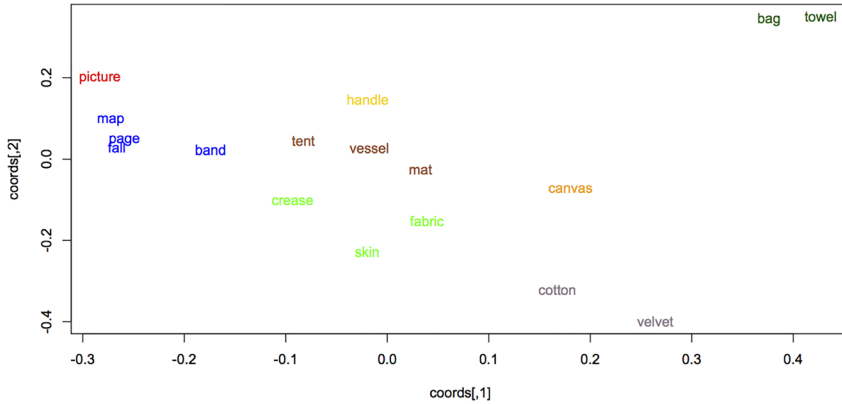
- (14) [...] find the joy in breaking your routine, Mellan suggests.
- (15) Steven's then-fiance, Traci Greer, broke their engagement and married Walter Millbank.

To sum up, while it seems that the most prototypical senses of the verb *break* are found in both constructions, less prototypical uses are restricted to one construction or the other. Typically, we observe that elements that cannot be acted upon, such as *dawn*, are only found with the INCCx whereas things that do not seem to have the ability to break without an external cause, i.e., that have to be acted upon in order to break, are restricted to the TCCx.

### 3.2.2 Tear

The verb *tear* is less polysemous and less frequent than *break* in these constructions but presents some interesting specificities too. After manually annotating 1,600 instances of *tear*, I identified 22 instances of *tear* in the INCCx and 70 in the TCCx for a total of 92 instances retained. The majority of instances extracted from the COCA were, unsurprisingly, instances of the resultative construction (including also phrasal verbs such as *tear apart*). There were also many instances of the setting construction such as *She tore her ACL playing golf*. Within the instances of the INCCx and the TCCx I identified 66 different themes, 20 with the INCCx and 51 with the TCCx for a total overlap of 7.58%.

The maps in Figures 6 and 7 show the difference in the number of different themes found with each construction but also the three main clusters of themes found with *tear*, namely FABRIC, PAPER and FLESH (such as *skin*, *cartilage* and *tendon*) which are actually the only ones found to be shared by the two constructions. As shown in Table 4, the main and only concrete meaning of *tear*, which is similar to *rip* is the only verb sense that is found with both constructions. However, it also appears that while this sense is shared by the two constructions, it is only found with themes that belong to the FABRIC, PAPER and FLESH families. While the INCCx is limited to these three types of themes, it is clear that the TCCx is used with a wider variety of themes such as *sandwich*, as illustrated in (16). This shows that in order for a theme to be used with *tear* in the INCCx, it must be construed as likely to tear on its own or at least as having properties that may facilitate the tearing event. With the TCCx, the theme is not as restricted as to its 'tearability' but rather has to be



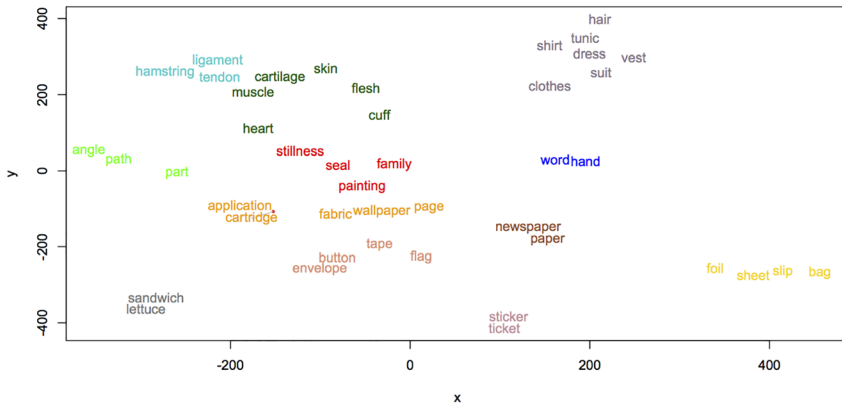
**Figure 6:** A semantic map of the themes that occur with *tear* in the INCCx.

something that can be acted upon. This further explains why the TCCx is also found with figurative senses.

- (16) She stood on her chair, making small grunting noises as she tore the sandwich and squeezed the bread and peanut butter into paste.

As can be seen in Table 4, two other senses of *tear* were identified with the TCCx and both are figurative uses of *tear*, as illustrated in (17) and (18).

- (17) I’ve seen how it has torn my family when you don’t accept that people are going to fall in love with who they fall in love with.



**Figure 7:** A semantic map of the themes that occur with *tear* in the TCCx.

**Table 4:** Shared and construction-specific verb senses of *tear*.

<i>tear</i>		
Intransitive non-causative construction	Shared	Transitive causative construction
	1. 1.1 FABRIC, 1.2 PAPER, & 1.3 FLESH	1.4. other themes 2. “cause to break apart (figuratively)” 3. “interrupt suddenly”

(18) [...] but a cough tore the stillness.

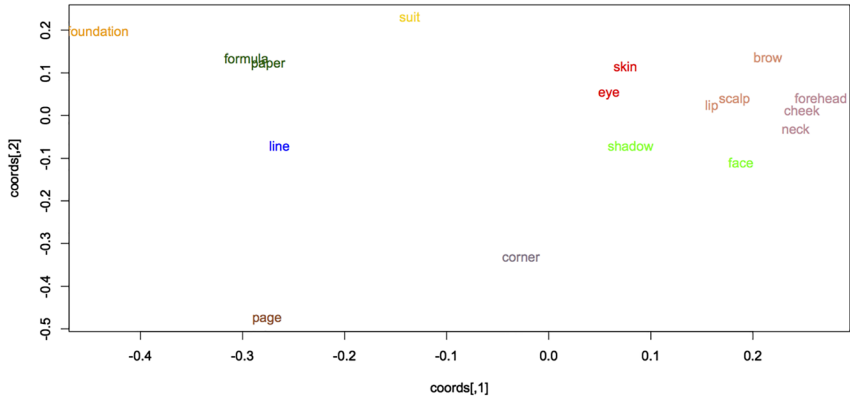
Crucially we see that only the most literal sense of *tear* is shared across constructions, and only with a limited number of themes.

### 3.2.3 Crease

Out of the 850 instances of *crease* extracted from COCA, I identified 112 instances of the INCCx and 247 instances of the TCCx, for a total of 359 instances retained. Interestingly, while my dataset contains substantially more instances of *crease* than *tear*, *crease* is clearly less varied as only 22 themes were found with the INCCx and 71 with the TCCx. This is explained by the fact that the meaning of *crease* is very specific and limited to one verb sense. This monosemy may also explain the unusually high amount of overlap of themes between the two constructions with 16.25% of themes shared.

As is visible from Figure 8, a majority of the themes that occur with *crease* in the INCCx relate to *SKIN* (cf. top right corner of the map). There are also themes that belong to the *PAPER* and *FABRIC* families. These three families are the only ones that are shared by the two constructions, as shown in Table 5. The themes *foundation* and *formula* both refer to creamy or liquid make-up products, as shown in (19) and (20).

- (19) Beverly Hills Mineral Foundation, \$65, is a mineral-based liquid formula that won't crease or fade.
- (20) Blot your face to get rid of excess oil and blend in foundation that has creased using a clean makeup sponge in a circular motion.



**Figure 8:** A semantic map of the themes that occur with *crease* in the INCCx.

Aside from the obvious larger number of themes found with the TCCx in Figure 9, we also find a whole cluster of themes that refer to types of SURFACE that various things can cause to crease, as in (21) and (22).

(21) Six flashing oars creased the surface of the river in synchronized strokes [...]

(22) [...] the highway only barely creasing the snowed-over sagebrush flat land.

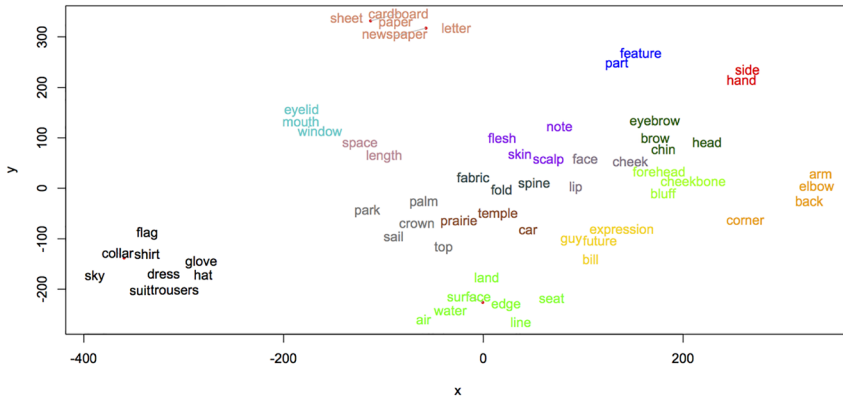
The distribution of *crease* across the two constructions seems to confirm our previous comments as to the constraints on the theme in each construction. Items such as FABRIC, PAPER and SKIN which seem to have certain properties that makes them very likely to crease but also likely to be creased by an external cause are shared by the two constructions. Interestingly, the group that is restricted to the INCCx is made of things that one does not want to crease, i.e., various makeup products.

### 3.2.4 Crinkle

Out of the 600 instances of *crinkle* manually annotated, I retained 330 instances: 204 in the INCCx and 126 in the TCCx. *Crinkle* was found with 73 different themes, 57

**Table 5:** Shared and construction-specific themes that occur with *crease*.

<i>crease</i>		
Intransitive non-causative Cx	Shared	Transitive causative Cx
	FABRIC, PAPER, and SKIN	
LIQUID PRODUCTS		SURFACE



**Figure 9:** A semantic map of the themes that occur with *crease* in the TCCx.

with the INCCx and 36 with the TCCx for a theme overlap of 27.4%. Again, a monosemous verb displays an unusually high theme overlap. *Crinkle* is semantically similar to *crease*, as both denote some type of intricate folding and just like *crease*, the only themes that are shared between the two constructions are of the FABRIC, PAPER and SKIN families, as shown in Table 6. Also, like *crease*, themes that refer to a type of SURFACE are limited to the TCCx.

As opposed to *crease* though, there are actually more themes that occur in the INCCx than the TCCx, as visible from Figures 10 and 11. Again, with the TCCx, we find themes that refer to a type of SURFACE, as illustrated in (23).

- (23) [...] a light breeze crinkled the surface of an otherwise smooth and moonlit sea.

Again, themes that can easily be construed as being likely to *crinkle* because of their properties are more likely to be shared, or, for that matter, be found in the INCCx. On the contrary, the surface of the sea requires external intervention to be created.

**Table 6:** Shared and construction-specific themes that occur with *crinkle*.

<i>crinkle</i>		
Intransitive non-causative Cx	Shared	Transitive causative Cx
	FABRIC, PAPER, and SKIN	
		SURFACE

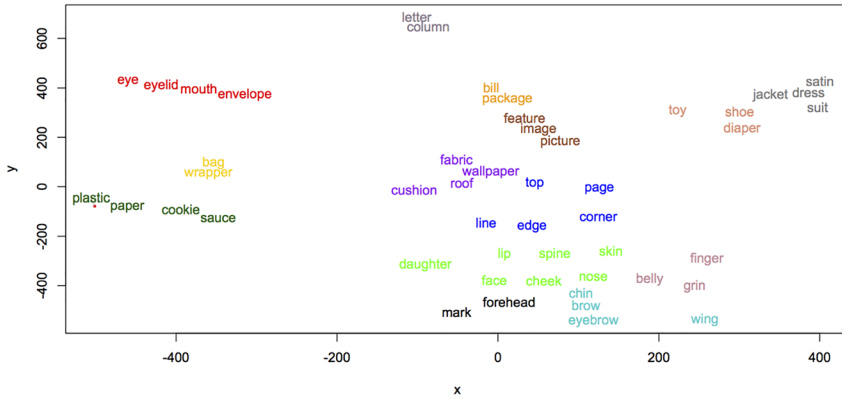


Figure 10: A semantic map of the themes that occur with *crinkle* in the INCCx.

### 3.2.5 Wrinkle

A total of 1,100 instances of *wrinkle* were extracted from the COCA and manually annotated. From these, 701 instances were retained: 215 instances of the INCCx and 489 instances of the TCCx. Despite the large number of instances found in the dataset, the variety of themes is somewhat limited, with only 87 different themes identified, 55 with the INCCx and 50 with the TCCx for a theme overlap of 20.69%.

The two maps in Figures 12 and 13 show that three clusters of themes are shared by the two constructions: *FABRIC*, *SKIN* and *SURFACE*. In this case, a small cluster of themes is restricted to the INCCx: *PEOPLE*. As may appear evident from the

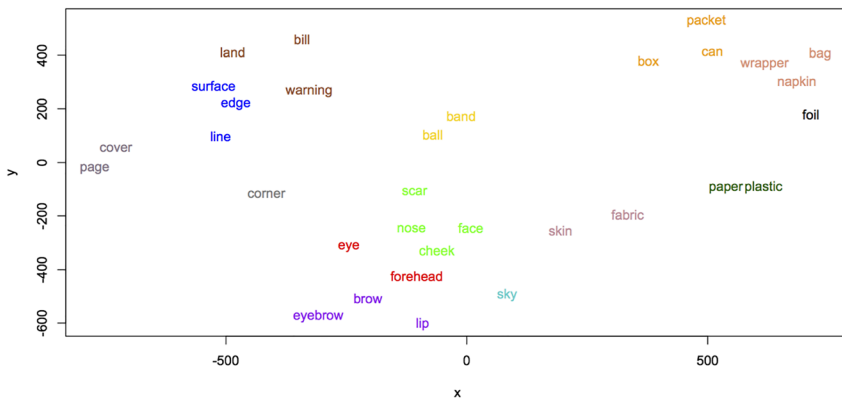
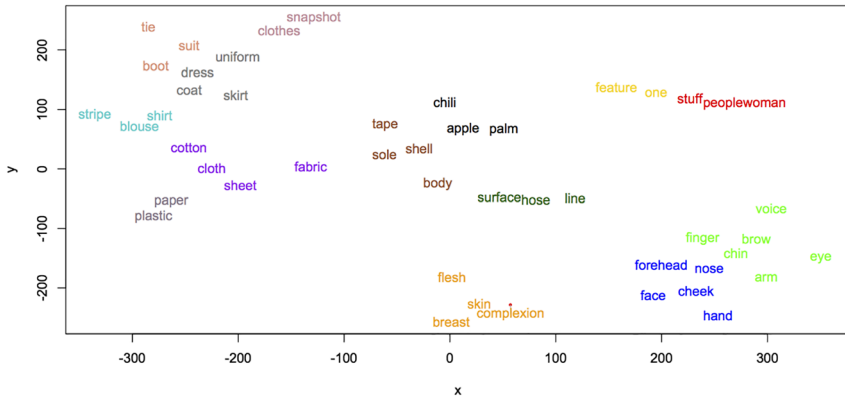


Figure 11: A semantic map of the themes that occur with *crinkle* in the TCCx.





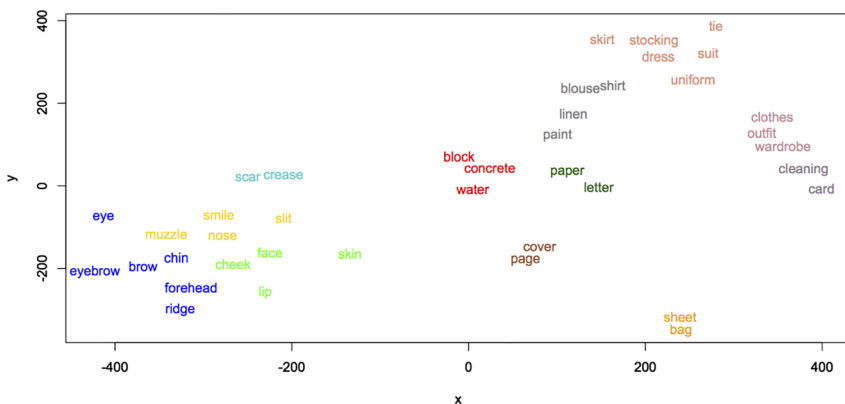
**Figure 12:** A semantic map of the themes that occur with *wrinkle* in the INCCx.

examples in (24) and (25), people can wrinkle (metonymically at least) but it seems unlikely that something or someone can be construed as wrinkling people. Note, however, that certain things can wrinkle people's skin.

(24) Once people start to stoop and wrinkle they all look alike to me.

(25) I didn't know they could prove I was wrinkling.

Our analysis of *wrinkle* also leads us to conclude that constraints on the use of these two constructions go not only beyond the level of the verb, but also beyond individual verb senses (see Table 7).



**Figure 13:** A semantic map of the themes that occur with *wrinkle* in the TCCx.

**Table 7:** Shared and construction-specific themes that occur with *wrinkle*.

<i>wrinkle</i>		
Intransitive non-causative Cx	Shared	Transitive causative Cx
FABRIC, SKIN, and SURFACE		
PEOPLE		

By means of conclusion to this section, it can be argued that a finer-grained analysis that includes each verb's themes and thus various senses reveals a certain amount of discrepancy between the two constructions. The most prototypical verb senses tend to be shared while less central verb senses tend to be restricted to one construction. At an even lower level of generalisation, we see that even when a verb sense is shared, it is regularly the case that some of the themes that correspond to this verb sense are actually restricted to one construction.

## 4 Discussion: putting the argument back in argument structure constructions

As William Croft states: “language once again successfully resists the attempts of linguists to make it neat and clean” (Croft 2003: 50). It seems clear at this point that so-called alternating constructions are in fact rather limited in their ability to alternate. As the analyses presented here show, the INCCx and the TCCx differ and this is observable at a very fine level of granularity. The identification of multiple verb senses in this analysis allowed a qualitative measure of the amount of information shared by the two constructions; as I have shown that only some of these verbs' senses are actually shared by the two constructions, thus following the observations made by Gilquin (2013) and Bernolet and Coleman (2016). Nevertheless, despite what more traditional verb-centred approaches argue, knowing a verb's meaning is not sufficient to predict its ability to occur in one construction, the other or both. Furthermore, constructional meaning (at least for these two constructions) cannot be described accurately when the focus is solely on the verbs used in these constructions. Rather, each argument structure construction exhibits features that constrain their use with certain theme-verb combinations.

Goldberg (2019: 59–60) argues that argument slots are constrained by the verb's semantics and that “because almost any action can be performed by various actors, and the vast majority of actions can be performed on various types of

entities, noun phrase arguments are much less particular than predicates when it comes to ASCs.” While this may be true to some extent, I would rather argue that by overlooking the interaction of the theme, verb and construction, one misses an important part of the picture. The method in this paper follows Langacker’s approach to semantic roles (Langacker 1991: 284) in that it looks for generalisations at the level of interaction with individual verbs (and construction) and draws more schematic descriptions of these roles from the individual instances. An extra layer of analysis is even added here as I have not only looked at theme-verb interaction but also at clusters of semantically related themes used with each verb in each construction. These clusters sometimes signal different verb senses (cf. *break* and *tear*) or show that even when the verb meaning remains constant across most uses of the verb, certain kinds of themes are restricted to one construction (cf. *crease*, *crinkle* and *wrinkle*). Through this careful investigation of themes, I have managed to identify certain constraints on the theme slot in each construction, namely that the theme in INCCx needs to have features that facilitate the event denoted by the verb, cf. for example themes of the PAPER family which are likely to tear, crease and crinkle. On the other hand, themes that occur in the TCCx need to be construed as being likely to be manipulated by a cause in the way denoted by the verb (e.g., SURFACE with *crease* and *crinkle*). I thus argue that in the case of the causative alternation at least, schematic constructional meaning is not to be identified by the verbs that prototypically occur in these constructions but rather by the interaction of the semantic role of the theme, the verb and the construction.

This is particularly important for cognitive approaches to language which defend a usage-based approach where speakers acquire their language from exposure to it. It is reasonable to assume that speakers pick up this type of low-level interactions and that these low-level generalisations guide their conceptualisation of events and the linguistic structures they choose to describe them. The role played by arguments in certain argument structure constructions has too often been overlooked. My final conclusion is thus that it is crucial that we put the argument back into argument structure constructions.

## Data availability statement

The data used for the creation of the semantic maps, based on the model by Perek (2021) is available on OSF: <https://osf.io/n324f/>. I have uploaded the annotated dataset for the five verbs presented here and the R script used to create the semantic maps on OSF: <https://osf.io/gcjrjv/>.

**Acknowledgements:** I would like to thank Florent Perek for helping me with the vector semantics analysis and for letting me use his model of distributional semantics, my PhD advisor, Maarten Lemmens for his help in developing this work, and Holger Diessel, who agreed to be my mentor for this special issue, for his very helpful comments on the first version of this paper. I would also like to thank two anonymous reviewers for their precious comments.

## References

- Bernolet, Sarah & Timothy Colleman. 2016. Sense-based and lexeme-based alternation biases in the Dutch dative alternation. In Jiyoung Yoon & Stefan Th. Gries (eds.), *Corpus-based approaches to construction grammar*. Amsterdam: John Benjamins.
- Boas, Hans C. 2003. *A constructional approach to resultatives*. Stanford: CSLI Publications.
- Bois, John W. Du. 1985. Competing motivations. In John Haiman (ed.), *Iconicity in syntax*, 343–365. Amsterdam: John Benjamins.
- Croft, William. 1991. *Syntactic categories and grammatical relations: the cognitive organization of information*. Chicago: Chicago University Press.
- Croft, William. 1998. Event structure in argument linking. In Miriam Butt & Wilhelm Geuder (eds.), *The projection of arguments: Lexical and compositional factors*, 21–63. Stanford: Center for the Study of Language and Information.
- Croft, William. 2003. Lexical rules vs. constructions: A false dichotomy. In Hubert Cuyckens, Thomas Berg, René Dirven & Klaus-Uwe Panther (eds.), *Motivation in language*, 49–68. Amsterdam: John Benjamins.
- Croft, William. 2012. *Verbs. Aspect and causal structure*. Oxford: Oxford University Press.
- Croft, William & D. Alan Cruse. 2004. *Cognitive linguistics*. New York: Cambridge University Press.
- Cruse, D. Alan. 1973. Some thoughts on agentivity. *Journal of Linguistics* 9. 11–23.
- Davidse, Kristin. 1991. *Categories of experiential grammar*. K.U. Leuven Unpublished PhD thesis.
- Davidse, Kristin. 1992. Transitivity/Ergativity: The janus-headed grammar of actions and events. In Martin Davis & Louise Ravelli (eds.), *Advances in systemic linguistics*, 105–135. London: Pinter.
- Delancey, Scott. 1984. Notes on agentivity and causation. *Studies in Language* 8. 181–213.
- Diessel, Holger. 2019. *The grammar network: How linguistic structure is shaped by language use*. Cambridge: Cambridge University Press.
- Dowty, David. 1991. Thematic proto-roles and argument selection. *Language* 67(3). 547–619.
- Evans, Vyvyan. 2009. *How words mean*. Oxford: Oxford University Press.
- Evans, Vyvyan. 2018. Conceptual vs. inter-lexical polysemy. In Lucy Pickering & Vyvyan Evans (eds.), *Language learning, discourse and cognition*. Amsterdam: John Benjamins.
- Faulhaber, Susen. 2011. Idiosyncrasy in verb valency pattern. *Zeitschrift für Anglistik und Amerikanistik* 59(4). 331–346.
- Fillmore, Charles J. 1977. The case for case reopened. In Peter Cole & Jerrold M. Sadock (eds.), *Grammatical relations*. New York: Academic Press.
- Firth, John. 1957. A synopsis of linguistic theory 1930–1955. In *Studies in Linguistic Analysis (Special volume of the philological society)*, 1–32. Oxford: Blackwell.

- Gilquin, Gaëtanelle. 2013. Making sense of collocation analysis. On the interplay between verb senses and constructions. *Constructions and Frames* 5(2). 119–142.
- Goldberg, Adele E.. 1995. *Constructions. A construction grammar approach to argument structure constructions*. Chicago: University of Chicago Press.
- Goldberg, Adele E. 2001. Patient arguments of causative constructions can be omitted: The role of information structure in argument distribution. *Language Sciences* 23. 503–524.
- Goldberg, Adele E. 2002. Surface generalizations: An alternative to alternations. *Cognitive Linguistics* 13(4). 327–356.
- Goldberg, Adele E. 2019. *Explain me this. Creativity, competition, and the partial productivity of constructions*. Princeton and Oxford: Princeton University Press.
- Gries, Stefan Th. & Anatol Stefanowitsch. 2010. Cluster analysis and the identification of collocational classes. In Sally Rice & John Newman (eds.), *Empirical and experimental methods in cognitive/functional research*, 73–90. Stanford: CSLI Publications.
- Gries, Stefan Th. & Dagmar Divjak. 2009. Behavioral profiles. A corpus-based approach to cognitive semantic analysis. In Vyvyan Evans & Stéphanie Pourcel (eds.), *New directions in Cognitive Linguistics*, vol. 24, 57–75. Amsterdam: John Benjamins.
- Gries, Stefan Th. & Anatol Stefanowitsch. 2004. Extending collocation analysis. A corpus-based perspective on ‘alternations’. *International Journal of Corpus Linguistics* 9(1). 97–129.
- Guerrero-Medina, Pilar. 2014. A lexico-paradigmatic approach to English setting constructions. In María de los Ángeles Gómez González, Francisco José Ruiz de Mendoza Ibáñez, Francisco González-García & Angela Downing (eds.), *The functional perspective on language and discourse: Applications and implications*, 133–148. Amsterdam: John Benjamins.
- Halliday, Michael Alexander Kirkwood. 1967a. Notes on transitivity and theme in English 1. *Journal of Linguistics* 3. 37–81.
- Halliday, Michael Alexander Kirkwood. 1967b. Notes on transitivity and theme in English 2. *Journal of Linguistics* 3. 199–244.
- Hampe, Beate & Doris Schönefeld. 2006. Syntactic leaps or lexical creation? More on ‘creative syntax’. In Stefan Th. Gries & Anatol Stefanowitsch (eds.), *Corpora in cognitive linguistics*, 127–157. Berlin: Mouton De Gruyter.
- Herbst, Thomas. 2011. The status of generalizations: Valency and argument structure constructions. *Zeitschrift für Anglistik und Amerikanistik* 59(4). 347–367.
- Herbst, Thomas. 2018. Is language a collocation? A proposal for looking at collocations, valency, argument structure and other constructions. In Pastor Cantos-Gómez & Moisés Almela Sánchez (eds.), *Lexical collocational analysis: Advances and applications*, 1–22. Cham: Springer.
- Hilpert, Martin & Florent Perek. 2015. Meaning change in a petri dish: Constructions, semantic vector spaces, and motion charts. *Linguistics Vanguard* 1(1). 339–350.
- Langacker, Ronald W. 1991. *Foundations of cognitive grammar*. (Vol. Volume 2: Descriptive Application). Stanford, CA: Stanford University Press.
- Langacker, Ronald W. 2008. *Cognitive grammar: A basic introduction*. Oxford: Oxford University Press.
- Lemmens, Maarten. 1998. *Lexical perspectives on transitivity and ergativity. Causative constructions in English*. Amsterdam: John Benjamins.

- Lemmens, Maarten. 2005. *De la sémantique lexicale à la typologie sémantique. Vers une sémantique basée sur l'activité langagière*. Lille: Université de Lille.
- Lemmens, Maarten. 2006. More on objectless transitives. *Constructions* 1. <http://journals.linguisticsociety.org/elanguage/constructions/article/view/2821.html>.
- Lemmens, Maarten. 2021. *New directions in cognitive linguistics. Usage-based perspectives on lexical and constructional semantics*. Shanghai: Foreign Language Education Press.
- Lenci, Alessandro. 2008. Distributional semantics in linguistic and cognitive research. *Rivista di linguistica* 20(1). 1–31.
- Levin, Beth. 1993. *English verb classes and alternations*. Chicago: Chicago University Press.
- Levin, Beth & Malka Rappaport-Hovav. 2005. *Argument realization*. Cambridge: Cambridge University Press.
- Levshina, Natalia & Kris Heylen. 2014. A radically data-driven construction grammar: Experiments with Dutch causative constructions. In Ronny Boogaart, Timothy Coleman & Gijbert Rutten (eds.), *Extending the scope of construction grammar*, 17–46. Berlin and Boston: De Gruyter Mouton.
- Miller, George A. & Walter G. Charles. 1991. Contextual correlates of semantic similarity. *Language & Cognitive Processes* 6(1). 1–28.
- OED Online. 2018. Oxford University Press.
- Padó, Sebastian & Mirella Lapata. 2003. Constructing semantic space models from parsed corpora. In *Proceedings of the 41st conference of the association for computational linguistics*. 128–135. Sapporo, Japan: Association for Computational Linguistics.
- Peirsman, Yves, Dirk Geeraerts & Dirk Speelman. 2010. The automatic identification of lexical variations between language varieties. *Natural Language Engineering* 16(4). 469–491.
- Perek, Florent. 2014. Rethinking constructional polysemy: the case of the English conative construction. In Dylan Glynn & Justyna A. Robinson (eds.), *Corpus methods for semantics. Quantitative studies in polysemy and synonymy*, 61–85. Amsterdam: John Benjamins.
- Perek, Florent. 2016a. Recent change in the productivity and schematicity of the way-construction: a distributional semantics analysis. *Corpus Linguistics and Linguistic Theory* 14(1). 149–188.
- Perek, Florent. 2016b. Using distributional semantics to study syntactic productivity in diachrony: A case study. *Linguistics* 54(1). 149–188.
- Perek, Florent. 2021. Distributional semantic models for english verbs and nouns. Available at: <https://doi.org/10.17605/OSF.IO/N324F>.
- Princeton University. 2010. About WordNet. Available at: <http://wordnet.princeton.edu>.
- R-Core-Team. 2013. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
- Sahlgren, Magnus. 2008. The distributional hypothesis. *Rivista di linguistica* 20(1). 33–53.
- Salton, Gerard, Andrew Wong & Chungshu Yang. 1975. A vector space model for automatic indexing. *Communications of the ACM* 18(11). 613–620.
- Schütze, Hinrich. 1992. Dimensions of meaning. In Robert Werner (ed.), *Proceedings of the 1992 ACM/IEEE conference on supercomputing*, 787–796. Los Alamitos, CA: IEEE Computer Society Press.
- Stefanowitsch, Anatol & Stefan Th. Gries. 2003. Collostructions: investigating the interaction between words and constructions. *International Journal of Corpus Linguistics* 8(2). 209–243.
- Talmy, Leonard. 2000. *Toward a cognitive semantics*, vol. 1. Cambridge, MA: MIT Press.
- Turney, Peter D. & Patrick. Pantel. 2010. From frequency to meaning: Vector space models of semantics. *Journal of Artificial Intelligence Research* 37. 141–188.

- Wild, Fridolin. 2007. An LSA package for R. In Fridolin Wild, Marco Kalz, Jan Van Bruggen & Rob Koper (eds.), *Mini-proceedings of the 1st European workshop on latent semantic analysis in technology-enhanced learning*, 11–12. Netherlands: Open University.
- Yi, Eunhyung, Jean-Pierre Koenig & Douglas Roland. 2019. Semantic similarity to high-frequency verbs affects syntactic frame selection. *Cognitive Linguistics* 30(3). 601–628.
- Yoshimura, Kimihiro & John R. Taylor. 2004. What makes a good middle? The role of qualia in the interpretation and acceptability of middle expressions in English. *English Language and Linguistics* 8. 293–321.