



**HAL**  
open science

## ”It is important to note that” partially productive patterns may count as constructions

Guillaume Desagulier

### ► To cite this version:

Guillaume Desagulier. ”It is important to note that” partially productive patterns may count as constructions. Jacques François. L’expansion pluridisciplinaire des grammaires de constructions, Presses universitaires de Caen, 2021, Bibliothèque de Syntaxe & Sémantique, 2381850015. halshs-01871034

**HAL Id: halshs-01871034**

**<https://shs.hal.science/halshs-01871034v1>**

Submitted on 10 Sep 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*It is important to note that*  
partially productive patterns may count as  
constructions

Guillaume Desagulier<sup>1</sup>

<sup>1</sup>MoDyCo — Université Paris 8, CNRS, Université Paris Nanterre, Institut  
Universitaire de France

**Abstract**

Determining what counts as a construction is a major bone of contention between redundant and non-redundant construction grammar taxinomies. Non-redundant taxonomic construction-grammar models posit that only maximally productive patterns qualify as constructions because they license an infinity of expressions. Redundant models claim that, despite subregularities and exceptions, partially productive patterns also count as constructions. I demonstrate that even patterns that are not fully productive at the most schematic level often have subregularities that are. I assess the productivity of a multiple-slot construction in the British National Corpus (XML edition): *it BE ADJ to V<sub>INF</sub> that*. The assessment involves hapax-based productivity measures, vocabulary growth curves, and LNRE models. I show that although the productivity of *it BE A to V<sub>INF</sub> that* is limited at its most schematic level, some partially filled subschemas such as *it BE hard/important/easy/difficult/possible/necessary/reasonable/impossible to V that* and *it BE ADJ to think/say/suggest/know/assume/realize/see that* are arguably productive.

Keywords: construction grammar, collocation, stance, productivity

## 1 Introduction

Far from being a unified framework, construction grammar is a family of approaches to language that is characterized by a set of shared tenets but also major differences. The main commonalities are the following: (a) constructions are form-meaning pairing (where form spans across morphology, syntax, phonology, etc.), (b) grammar is a structured inventory of such constructions, and (c) this inventory is non-derivational and monostratal.

Differences pertain to the nature and contents of the inventory. Cognitive Construction Grammar (henceforth CCG) is non-reductionist because it allows for redundancies and exceptions (Goldberg 1995, 2003, 2006, 2009). In other words, any form-meaning pairing can be claimed to be a construction as long as its overall meaning is not strictly compositional. Conversely, Berkeley Construction Grammar (henceforth BCG) is reductionist (Fillmore, Kay, & O'Connor 1988; Fillmore 1997; Kay & Fillmore 1999). Only schematic, general, and maximally productive patterns count as constructions.

For example Kay considers that the multiple-slot construction *A as NP* is not a construction but a mere pattern of coining because the pattern cannot be used to understand or freely create new expressions of the same kind. Yet, a corpus study reveals that even though *A as NP* may not be fully productive at the schematic level, it generates subschemas whose productivity is doubtless (Desagulier 2015).

In this paper, I further argue against an introspective approach to productivity in Construction Grammar and reassert that even patterns that are not fully productive at the most schematic level often have subregularities that are. I assess the productivity of *it BE ADJ to V<sub>INF</sub> that* in the British National Corpus (XML Edition). This multiple-slot pattern is exemplified below:

- (1) **It is important to note that** 33% of the sound's total land area will be protected. (BNC–CRC)
- (2) **It is fair to say that** it has the sweep and scope of an encyclical. (BNC–CRK)
- (3) Even so, **it was hard to believe that** his birthday in two days' time would be only his twelfth. (BNC–FRF)

Presumably, *it BE ADJ to V<sub>INF</sub> that* is not likely to be granted the construction status in the BCG framework because the list of candidates to the adjectival and verbal slots is limited, like the number of different A–V combinations.

Given the radical claim that productivity sets the dividing line between a pattern of coining and a construction in BCG, productivity measures should be used to confirm introspection. In the wake of quantitative approaches in the fields of morphology (Baayen 1989; Baayen & Lieber 1991; Baayen 1993) and morphosyntax (Zeldes 2012), I adopt the idea that productivity is a factor of both a large number of low-frequency words and a low number of high-frequency words. This is captured by the correlation between the number of hapax legomena of a given grammatical category and the number of neologisms in that category. The productivity of the rule at work is a direct effect of this correlation.

I combine these hapax-based productivity measures with symmetric and asymmetric association measures, namely  $\chi^2$  and  $\Delta P$  (Allan 1980). Thanks to this combination, I show that although the productivity of *it BE ADJ to V<sub>INF</sub> that* is limited at its most schematic level, some partially filled subschemas such as *it BE hard/important/easy/difficult/... to V that* and *it BE ADJ to think/say/assume/... that* are arguably productive.

Section 2 outlines key issues regarding the empirical measure of productivity in a construction grammar framework. Section 3 presents the corpus and the data. Section 4 lists and discusses the methods used to assess constructional productivity. Section 5 reports the results, which are discussed in Section 6. In Section 7, I conclude by stressing the contribution of partially productive patterns to our understanding of language competence.

## 2 Key issues and previous works

Much to the surprise of many linguists working in the Construction Grammar framework (including BCG), Kay (2013) has turned BCG into a generative approach based on a strong distinction between *grammar* and the *meta-grammar*. Grammar consists of “the strictly linguistic information required to produce and interpret any set of expressions of the language *and no more*”. In this respect, patterns such as *let alone*, *what's X doing Y*, or *all-clefts* qualify as constructions because they license an infinity of expressions. The meta-grammar contains “patterns [that] are neither necessary nor sufficient to produce or interpret any set of expressions of the language [...]”. According to Kay, *A as NP*, illustrated below, is not a construction but a mere

“pattern of coining” because (a) knowing the pattern is not sufficient to licence and interpret new expressions, and (b) A–NP combinations are idiomatically constrained and speakers cannot freely coin new expressions.

- (4) ‘Come off it, mate,’ Bragg exclaimed, ‘it’s a bloody weapon! You could kill a man with it, **as easy as spit.**’ (BNC–ANL)
- (5) a. He was eighteen and **nervous as a long-tailed cat in a room full of rocking chairs.** (COCA–FIC–ChicagoRev)  
 b. I’m **nervous as a cat** when I’m away from Mamma. (COCA–FIC–Bk:DeadlyAffairAt)  
 c. David himself was as **nervous as a bird.** (COCA–FIC–SouthernRev)  
 d. I’m **nervous as a dog shitting razor blades.** (COCA–FIC–SouthernRev)  
 e. I’m **as nervous as a mouse.** (G. Bernard Shaw – *Arms and the Man*)  
 f. She was **nervous as a mouse in a cat-house** (...). (COCA–FIC–ArkansasRev)  
 g. I was **nervous as a mouse in a python cage** (...) (Jim D. Jordan – *Bone Digger*)  
 h. ?I am **nervous as a flea.**

Admittedly, the semantic relationship between *easy* and *spit* in (4) is far from obvious. Kay suggests that this might block a naive speaker from inferring the adjectival intensification conventionally associated with the pattern (‘very easy’). Undoubtedly, some *A as NP* expressions are idiomatic enough to constrain the range of NPs used as paragons to perform the intensifying function. For example, the examples in (5) suggest that NPs denoting animals are common in *nervous as NP*, but the list of candidates cannot be extended freely as the unacceptability of (5h) shows. However, speakers commonly amend a well-entrenched subschema at the level of the NP postmodifier: compare (5a) to (5b), and (5e) to (5f) and (5g).

There are two main problems with deciding upon the constructional status of a pattern based on maximal productivity at the most schematic level. First, arbitrary sets of bigrams may well be very productive according to any of the hapax-based productivity measures, but this does not mean that they constitute a construction (Baayen 2001, 221). Second, even the most schematic construction will not generate an infinity of instances. For example, the *V someone A* construction (Boas 2003) is highly schematic. However, as Bybee (2010) observes, the two slots are not equally productive. Indeed, the range of verbs is quite limited (*make, send, drive*), unlike the range of adjectives, which is relatively open ended (*crazy, nuts, sad, mad, angry, etc.*).<sup>1</sup> This means that intermediate productivity levels must be recognized at the subschematic level.

In Desagulier (2015), I use a combination of hapax-based productivity measures and association measures to show that *A as NP* is undeniably productive at the subschematic level. These productive subschemas may function as productive intermediate schemas for more specific instantiations of the construction. Among the most productive subschemas, we find the following: *black/white/big/bright/cold/... as NP* and *A as hell*.

This paper is a follow-up to the abovementioned study. I focus on a different multiple-slot construction: *it BE ADJ to V<sub>INF</sub> that*. It is exemplified in (6)–(9).

- (6) **It is also quite untrue to say that** any foreigners obtained land. (BNC–HH3)  
 (7) **It is tempting to assume that** the same rule applies,... (BNC–B7H)  
 (8) ... and **it is pleasing to note that** many more are in the pipeline. (BNC–A11)  
 (9) ... but **it is misleading to suggest that** no incentives to good performance exist except market forces. (BNC–G1C)

<sup>1</sup>Of course, selectional restrictions apply depending on the verb. The range of adjectives is wider with *make* than with *drive*.

On the one hand, *it BE ADJ to V<sub>INF</sub> that* meets the formal criteria for being a considered a construction. It has an identifiable syntax, namely a *to*-complement clause controlled by a stance adjective, and it performs a specific function, namely stance marking. Simply defined, stance has to do with the expression of a position with respect to the form or the content of an utterance.<sup>2</sup> On the other hand, the pattern is likely to be considered as non productive by Kay (2013) because (a) the list of candidates to the adjective and verb slots is limited, and (b) the number of adjective-verb combinations is not open-ended. The purpose of this paper is therefore to show that, although we have good reasons to believe that the construction is a priori unproductive, areas of productivity can be found at the subschematic level.

Stance has been studied as a concept (Martin & White 2005; Englebretson 2007; Hunston 2011), alongside attitude, appraisal, evaluation, modality, viewpoint, intersubjectivity, etc., in which case it has been approached in a qualitative way (Martin & White 2005; Englebretson 2007). It has also been studied as a lexico-grammatical phenomenon (Biber 2006), which lends itself to a quantitative, corpus-based approach (Biber 2006; Hunston 2007, 2011).

Stance marking can be overt, when stance attribution is explicit, or covert, when stance attribution is implicit. When stance is covert, it is often difficult to decide whether it stems from the speaker/writer. Following Biber et al. (1999, 976–978), Gray and Biber (2014, 222) place stance markers “along a cline ranging from explicit to implicit to ambiguous attribution to the speaker/writer”. The most explicit attribution of stance to the speaker/writer is characterized by constructions that contain a pronoun, such as verb/adjective-controlled complement clauses with first person subjects (10), extraposed verb/adjective-controlled complement clauses with *me/us/to me/to us* (11), or *my/our* + N-controlled complement clauses (12).

(10) **I guess** we’re both losers, Mark. (BNC–AC2)

(11) **It seems to us quite unrealistic** to say that he would have felt insulted. (BNC–GVR)

(12) **Our desire to** please them will take precedence over our own needs (. . .). (BNC–CEF)

The least explicit attribution of stance to the speaker/writer is characterized by N-controlled complement clauses (13) and N + PP constructions (14).

(13) **The claim that** literary criticism can influence life, if you let it, or even transform it, is not in itself absurd. (BNC–CKN)

(14) This was due to the low frequency of the susceptibility allele (4310-5) and **the resulting low probability that** any of the affected patients were homozygous. (BNC–CNA)

The *it BE ADJ to V<sub>INF</sub> that* construction stands halfway between explicit and covert stance attribution, along with constructions that consist of (semi-)modals (15), adverbials (16), and probability verbs that control complement clauses (17).

(15) The lexical items in a taxonomy **may** be thought of as corresponding to classes of things in the extra-linguistic world. (BNC–FAC)

(16) **Admittedly** not everyone was impressed. (BNC–EAY)

(17) Good Englishness doesn’t dress up, indeed it **tends to** hide itself. (BNC–A0U)

According to Gray and Biber (2014), the lexico-grammatical approach hinges on “a set of lexical items that indicate a particular evaluation, attitude/emotion, or degree of epistemic certainty or doubt are taken to mark a speaker/writer’s stance when they occur in particular grammatical contexts [. . .].” The lexico-grammatical approach has given pride of place to individual

<sup>2</sup>In fact, stance is far more complex to grasp because “while individual researchers do tend to operationalize *stance* within their own work, definitions and understandings are not necessarily shared as common ground from one scholar to the next” (Englebretson 2007, 2).

lexemes such as pronouns (Myers & Lampropoulou 2012), adverbs (Diani 2008), verbs (Aijmer 2009), or interactional particles (Morita 2015), for example. Other studies have embraced morpho-syntactic or lexico-syntactic structures such as adverbials (Biber & Finegan 1988), the passive voice (Baratta 2009), *that*-clauses (Hyland & Tse 2005; Charles 2007), extraposed *it*-clauses (Hewings & Hewings 2002), or LE constructions in Mandarin (Chang 2009). These works assume that markers convey stance in themselves and/or appear frequently in contexts where stance is relevant.

I propose a shift from a lexico-grammatical to a CCG approach. At first sight, this approach is not fundamentally different from the abovementioned studies. The difference comes from the status that one grants to these stance-marking units. In CCG, “[a]ny linguistic pattern is recognized as a construction as long as some aspect of its form or function is not strictly predictable from its component parts or from other constructions recognized to exist” (Goldberg 2003). Previous works adopting a construction-grammar approach to overt stance marking include Dancygier (2012) and Biq (2004). Stance has also been studied in gesture (Dancygier & Sweetser 2012). I contend that the pattern *it* BE ADJ *to* V<sub>INF</sub> *that* serves as a prepackaged unit available to speakers in the context of stance marking. Although prepackaged, the construction is not fixed once and for all. It has two loci of variation: the adjective and the verb slots. I hypothesize that, although largely entrenched at the most schematic level, the construction may be productive at subschematic levels.

### 3 Corpus and data

As Gray and Biber (2014) point out, most studies of stance follow a comparative-register approach. Hewings and Hewings (2002) find that *it*-clauses perform four interpersonal functions: hedging, marking the writer’s attitude, emphasis, and attribution. More specifically, *it*-clauses are used by writers to persuade readers of the validity of their claims. The authors compare the uses of this construction in two corpora. One corpus is a collection of articles in the field of business studies. The other corpus is a collection of dissertations by MBA students who are non-native speakers of English. Hewings and Hewings find that the student writers tend to be more overtly persuasive than the academic writers.

Gray and Biber (2014) observe that extraposed complement clauses (specifically extraposed adj + *to/that*-clause) are frequent in academic writing because of the lack of explicit attribution of stance to the writer. In (18), we can easily assume that it is the academic who deems it reasonable to expect that a change in people’s attitudes and beliefs towards attempted suicide will be impactful. At the same time, because the origin of stance is left implicit by the *it* BE ADJ *to* V<sub>INF</sub> *that* construction, the writer’s belief is presented as widely shared.

- (18) It is reasonable to expect that public attitudes and beliefs about attempted suicide will affect its incidence. (BNC–B30)

Using a 14.3M-word sample from the corpus investigated in the Longman Spoken and Written English corpus, Gray and Biber compare the distributions of simple vs. extraposed adjectives followed by *to/that*-complement clauses across three registers in American English: academic prose (5.3M word tokens), newspapers (4.9M word tokens), and conversation (4.1M words). There are two main findings: (a) complement clauses controlled by a stance adjective are more common in academic prose than in the other two registers, and (b) extraposed clauses occur three to nine times more frequently in academic English than in news or conversation.

To verify if this tendency is verified in a corpus of British English, I extracted all instances of *it* BE ADJ *to* V<sub>INF</sub> *that* from the British National Corpus (XML Edition), which consists

of about 98+ million word tokens across 4049 corpus files. The query returned 2136 matches. Each adjective was manually annotated with one of the six semantic classes that Gray and Biber (2014, 248) found in the context of the specific construction:

- ability and willingness: e.g. *able, keen, prepared,*
- attitude and emotion: e.g. *surprising, sad, arrogant,*
- ease and difficulty: e.g. *difficult, hard, easy,*
- epistemic (certainty): e.g. *true, right,*
- epistemic (likelihood): e.g. *possible, plausible,*
- evaluation: e.g. *fair, important, good.*

The semantic classes were cross-tabulated with the eight text genres of the BNC:

- ACPROSE: academic writing,
- FICTION: published fiction,
- NEWS: news and journalism,
- NONAC: published non-fiction,
- UNPUB: unpublished writing,
- OTHERPUB: other published writing,
- CONVRSN: spoken conversation,
- OTHERSP: other spoken.

The results are listed in a contingency table (Table 1).

Table 1: The distribution of adjective classes by BNC text genres in the construction

	ACPROSE	CONVRSN	FICTION	NEWS	NONAC	OTHERPUB	OTHERSP	UNPUB
ability.willingness	1	0	0	0	8	7	1	0
attitude.emotion	29	0	14	12	38	29	1	5
ease.difficulty	112	0	51	20	115	56	0	9
epistemic.certainty	65	1	6	5	49	25	8	5
epistemic.likelihood	57	0	4	3	43	7	0	2
evaluation	533	0	24	48	450	215	13	65

Although modest in size, the contingency table is hard to synthesize with the naked eye. To facilitate exploration, Table 1 was submitted to correspondence analysis (for an overview, see Desagulier 2017, Section 10.4). Figure 1 displays its graphic output. Semantic classes are in blue, and text genres in magenta. Our results differ from Gray and Biber (2014)’s findings. Academic prose does not stand out with respect to news and conversation in the BNC. ACPROSE clusters in the lower left corner of the plot, along with UNPUB and NONAC. In this text genre, the construction shows a preference for the expression of evaluation and likelihood. In the NEWS genre, the construction shows a preference for attitude/emotion and ease/difficulty. The spoken text genres (CONVRSN and OTHERSP) stand out as they cluster far from the other text genres in the upper left corner. A quick glance at Table 1 shows that this is because the *it BE ADJ to V<sub>INF</sub> that* construction occurs only 24 times in spoken text genres. Unlike what Gray and Biber (2014) find in their corpus of American English, the *it BE ADJ to V<sub>INF</sub> that* is not particularly characteristic of academic writing in the British National Corpus.

Because the preferences of the *it BE ADJ to V<sub>INF</sub> that* construction are likely to differ based on the corpus that one explores, my main goal here is not to determine with respect to what register the construction is the most productive but (a) to assess its overall productivity based on a finite sample (a corpus), and (b) go beyond naive productivity assessments based on intuition only or inappropriate measures.

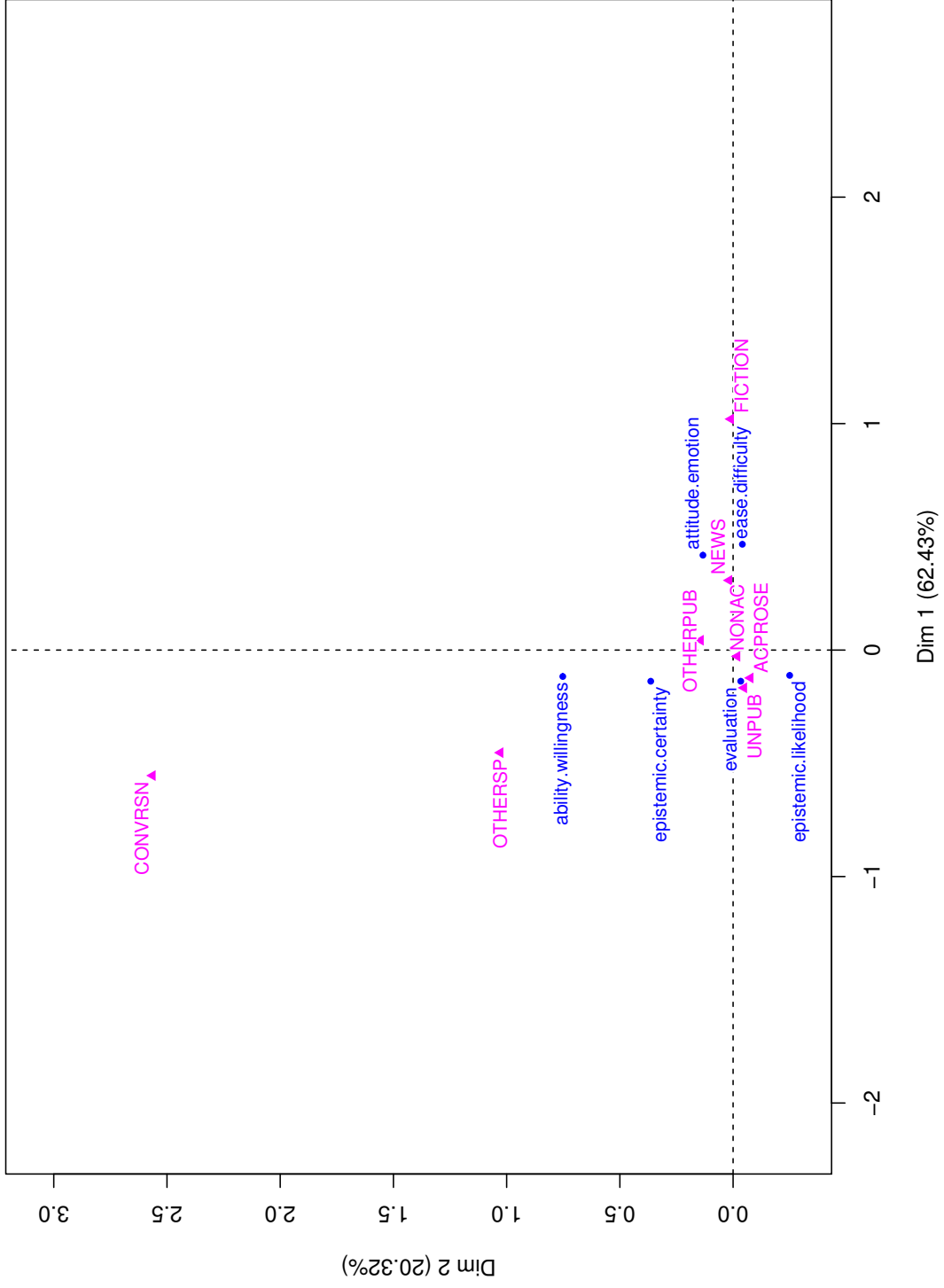


Figure 1: The distribution of adjective classes by BNC text genres (correspondence analysis)



## 4 Methods

Perhaps one of the most common measures used for the capture of productivity is type frequency. Baayen (1993) denotes it  $V(C, N)$ , where  $V$  stands for type, i.e. the frequency of a linguistic category  $C$  observed in a corpus of  $N$  tokens. There are three problems with  $V(C, N)$ . First, it does not account for skewed distributions in a corpus. Instead, it assumes that types are distributed evenly, which is hardly ever the case. Second, because it measures past/realized productivity (Baayen 2009, 902), it does not indicate whether a linguistic category has reached its peak of productivity or not. Third, it does not discriminate between established uses and innovating uses.

Suppose you want to compare the productivity of two novels by Herman Melville based on type, namely *Moby Dick* (173333 word tokens, 25728 word types) and *Bartleby, the Scrivener: A Story of Wall Street* (12507 word tokens, 3792 word types). In itself, the comparison is unfair because the former is a huge novel whereas the latter is a short story. According to  $V(C, N)$ , *Moby Dick* is far more productive than *Bartleby*. To keep track of vocabulary development across a given corpus, one can plot the number of types against the number of tokens at multiple intervals. One obtains a vocabulary growth curve (henceforth VGC) (Baayen 1993). At first, the curve is expected to rise steeply as most tokens define a new type. Then, as more text is scanned, more and more tokens are included in already defined types, and the curve flattens out gradually. The VGC in Figure 2 shows that the type productivity of *Moby Dick* (blue line) exceeds by far the type productivity of *Bartleby* (red line).

For a fair comparison, we should ask how lexically rich *Bartleby* would be if it were the same size in word tokens as *Moby Dick*. In order to answer this question, we have to be able to predict the vocabulary size of a larger corpus based on the vocabulary size of a smaller corpus. This can be done thanks to a simple yet crucial principle underlying productivity measures, as formulated by Baayen (1989), Baayen and Lieber (1991), and Baayen (1993) in morphology: the number of hapax legomena of a given schema correlates with the productivity of this schema. Once productivity is operationalized along these lines, a linguistic category is said to be productive when it is characterized by a Large Number of Rare Events. LNRE models (Baayen 2001) are designed to extrapolate based on the empirical distribution of a linguistic category and estimate its type count if we were to extend the size of the corpus.

To predict the vocabulary size of a larger corpus based on the vocabulary size of a smaller corpus, four steps are needed. First, one converts the data into a frequency spectrum, that is to say a table of frequencies of frequencies. Second, one fits a LNRE model to the spectrum object, choosing among three models:

- the Zipf-Mandelbrot model
- the finite Zipf-Mandelbrot model
- Generalized Inverse Gauss-Poisson model

Third, once a satisfactory model has been fitted to the frequency spectrum, values for the vocabulary size and the spectrum elements can be estimated within the range of the actual corpus size (interpolation) and beyond (extrapolation). Fourth, the results can be plotted on an enhanced vocabulary growth curve (VGC).

In Figure 3, the growth curves are extrapolated to twice the actual size in word tokens of *Moby Dick*. This graphic representation allows for a fair comparison in terms of productivity. If *Bartleby* were the same size as *Moby Dick*, it would not be as lexically rich because it would contain much fewer hapax legomena.

In morphology, a family of measures, based on Baayen (1989) and further described in Baayen and Lieber (1991) and Baayen (1993, 2001, 2009), *inter alia*, taps into the idea that the

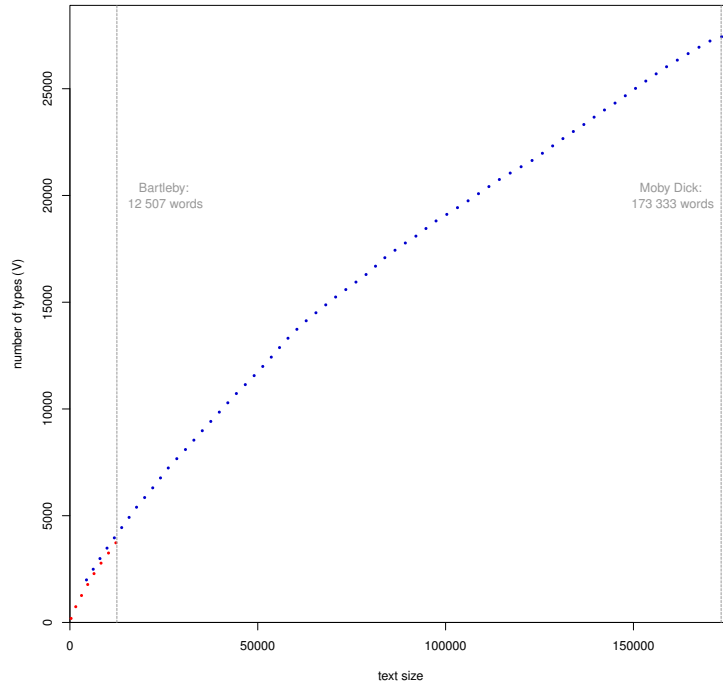


Figure 2: Vocabulary growth in Herman Melville's *Moby Dick* and *Bartleby, the Scrivener* (empirical)

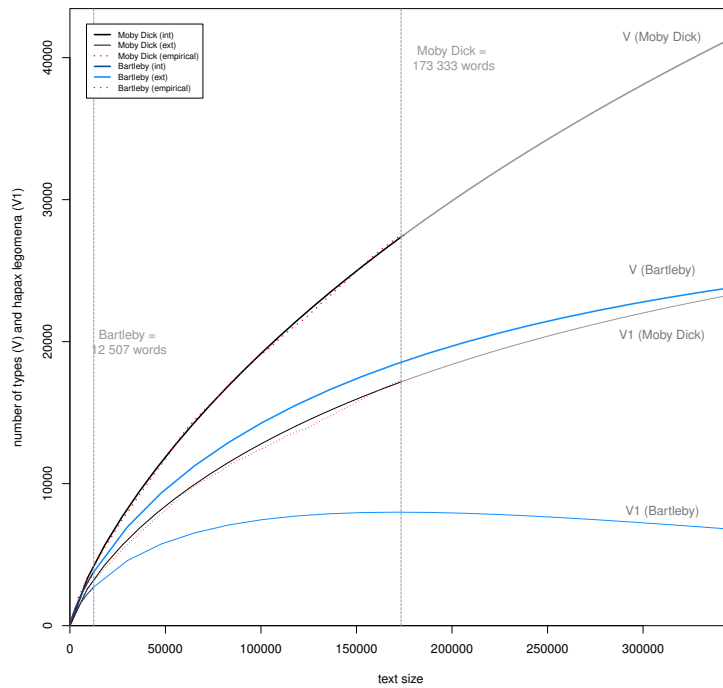


Figure 3: Vocabulary growth in Herman Melville's *Moby Dick* and *Bartleby, the Scrivener* (empirical, interpolated, and extrapolated)

productivity of a category correlates with the number of neologisms in this category. Expanding productivity is the hapax-conditioned degree of productivity. It is the ratio of the number of hapax legomena with a given affix and the sum of all hapax legomena in the corpus:

$$\mathcal{P}^* = \frac{V(1, C, N)}{V(1, N)}. \quad (\text{i})$$

Potential productivity ( $\mathcal{P}$ ) is another measure which, unlike  $\mathcal{P}^*$ , accounts for whether a productive morphological process is saturated.  $\mathcal{P}$  is the ratio of the number of hapax legomena with a given affix and the sum of all tokens that contain the affix:

$$\mathcal{P} = \frac{V(1, C, N)}{N(C)}. \quad (\text{ii})$$

$\mathcal{P}$  takes a value ranging between 0 and 1. The closer  $\mathcal{P}$  is to 1, the higher the probability of encountering new types, and the larger the productivity. Interestingly,  $\mathcal{P}$  is the slope of the tangent to a VGC at its endpoint. The steeper the tangent, the more productive the linguistic category. Because  $\mathcal{P}$  measures the probability of finding yet unseen types to the detriment of observed types, Baayen has devised a productivity measure that evaluates  $\mathcal{P}$  in the light of actual use. It is denoted  $P^*$  and it is the ratio of type frequency and  $\mathcal{P}$ :

$$P^* = \frac{V}{\mathcal{P}}. \quad (\text{iii})$$

Zeldes (2012) and Desagulier (2015) show that, although originally designed for the study of affixes in morphology, these hapax-based productivity measures apply equally well to the study of multiple-slot constructions. In the context of the *it* BE ADJ *to* V<sub>INF</sub> *that* construction, productivity can be measured at several levels:

- the concatenation of the adjective and the verb slots (e.g. *interesting–note*),
- the adjective slot (*interesting*) or the verb slot (*note*).

The first option is interesting because of the semantic interdependence between the adjective and the verb. Each fused type can therefore be tested for productivity. If a specific Adj–V<sub>INF</sub> concatenation has not been observed yet in the corpus, it is treated as a hapax legomenon. If it has, it leaves the hapax list. The second option is equally interesting because it captures the respective combinatorial possibilities of adjectives and verbs.

Because the *it* BE ADJ *to* V<sub>INF</sub> *that* construction has two slots, I will also measure the collocativity between the adjective and the verb by means of two association measures:  $\chi^2$  and  $\Delta P$  (Allan 1980). In the context of covarying-collexeme analysis (Gries & Stefanowitsch 2004; Stefanowitsch & Gries 2005), a method designed to measure collocation strength, i.e. the association between the two slots of a construction,  $\chi^2$  will serve to determine those Adj–V<sub>INF</sub> pairs whose association is highly conventionalized, autonomous, and non-analyzable. I hypothesize that the degree of association between the adjective and the verb can be used tentatively as an inverse measure of productivity: the higher the collocation score, the higher the degree of fixedness, and the lesser the productivity.

Unlike  $\chi^2$ , which does not say anything as to whether the lexeme in the first slot attracts the lexeme in the second slot more than vice versa,  $\Delta P$  is directional. Inspired by the rejection of classical conditioning to the benefit of associative learning (Rescorla 1968; Wagner & Rescorla 1972), Allan (1980) devised  $\Delta P$  as a one-way-dependency statistic to measure the asymmetric dependency between a cue and an outcome in behavioral experiments. As adapted by Ellis

(2006), Ellis and Ferreira-Junior (2009) and applied later by Gries (2013), Desagulier (2015) in the study of collocation,  $\Delta P$  takes a contingency table such as Table 2a as input. In the context of our construction, four configurations are possible depending on whether the adjective/verb is present/absent, as shown in Table 2b.

Table 2: Contingency tables for the computation of  $\Delta P$

(a) Contingency table involving a cue (C) and an outcome (O)	O: present    O: absent		(b) Contingency table involving an adjective and a verb	V: present    V: absent	
C: present	a	b	Adj: present	a	b
C: absent	c	d	Adj: absent	c	d

$\Delta P$  is calculated as follows:

$$\begin{aligned}\Delta P &= p(O|C) - p(O|\neg C) \\ &= \frac{a}{a+b} - \frac{c}{c+d}\end{aligned}\tag{iv}$$

The closer  $\Delta P$  is to 1, the more  $C$  increases the likelihood of  $O$ . By extension, when  $C$  increases the likelihood of  $O$ ,  $C$  is a good predictor of  $O$ . Conversely, the closer  $\Delta P$  is to  $-1$ , the more  $C$  decreases the likelihood of  $O$ . If  $\Delta P = 0$ , there is no covariation between  $C$  and  $O$ .

For the sake of illustration, let us apply Table 2b to one concrete example of *it* BE ADJ to  $V_{\text{INF}}$  *that*: *it is insane to imagine that*. Table 3 is the input contingency table showing the joint distribution of *insane* and *imagine*.

Table 3: Contingency table involving the adjective *insane* and the verb *imagine*

	<i>imagine</i> : present	<i>imagine</i> : absent
<i>insane</i> : present	1	0
<i>insane</i> : absent	42	2093

Two  $\Delta P$  values must be computed, depending on whether the cue is the adjective and the outcome is the verb, or vice versa:

$$\begin{aligned}\Delta P_{(\textit{imagine}|\textit{insane})} &= p(\textit{imagine}|\textit{insane}) - p(\textit{imagine}|\neg\textit{insane}) \\ &= \frac{a}{a+b} - \frac{c}{c+d} \\ &= \frac{1}{1+0} - \frac{42}{42+2093} \\ &= 0.9803279\end{aligned}\tag{v}$$

$$\begin{aligned}\Delta P_{(\textit{insane}|\textit{imagine})} &= p(\textit{insane}|\textit{imagine}) - p(\textit{insane}|\neg\textit{imagine}) \\ &= \frac{a}{a+c} - \frac{b}{b+d} \\ &= \frac{1}{1+42} - \frac{0}{0+2093} \\ &= 0.02325581\end{aligned}\tag{vi}$$

If  $\Delta P_{(imagine|insane)} - |\Delta P_{(insane|imagine)}|$  is positive, then *insane* is a better predictor of *imagine* than vice versa. Conversely, if  $\Delta P_{(imagine|insane)} - |\Delta P_{(insane|imagine)}|$  is negative, then *imagine* is a better predictor of *insane* than vice versa. If  $\Delta P_{(imagine|insane)} - |\Delta P_{(insane|imagine)}|$  is null, then no word is a good predictor of the other.

$$\begin{aligned} \Delta P_{\text{DIFF}} &= \Delta P_{(imagine|insane)} - |\Delta P_{(insane|imagine)}| \\ &= 0.9803279 \end{aligned} \tag{vii}$$

The  $\Delta P$  difference is positive, which implies that *insane* is a better predictor of *imagine*. This is unsurprising because *imagine* co-occurs with other adjectives such as *absurd*, *hard*, *extraordinary*, *implausible*, etc. On the other hand, *insane* co-occurs exclusively with *imagine* in the BNC. I hypothesize that good predictors are more productive than worse predictors. I therefore use  $\Delta P$  tentatively to unveil aspects of subschematic productivity that hapax-based measures may leave aside.

## 5 Results

Type, hapax legomena, and therefore productivity measures depend on how the construction is parsed. Four parsing levels are inspected:

- exact matches,
- Adj- $V_{\text{INF}}$ ,
- Adj slot,
- $V_{\text{INF}}$  slot.

Each level is exemplified in Table 4.

Table 4: Parsing levels for the productivity assessment of the construction

corpus file	exact matches	Adj- $V_{\text{INF}}$	Adj slot	$V_{\text{INF}}$ slot
A61.xml	<i>It was sad to think that</i>	sad_think	sad	think
AB4.xml	<i>it is sad to think that</i>	sad_think	sad	think
ASC.xml	<i>it is sad to think that</i>	sad_think	sad	think
C9E.xml	<i>It is sad to see that</i>	sad_see	sad	see
CAJ.xml	<i>It is sad to see that</i>	sad_see	sad	see
CHK.xml	<i>It is sad to see that</i>	sad_see	sad	see
CKX.xml	<i>it is sad to note that</i>	sad_note	sad	note
HHW.xml	<i>it is sad to reflect that</i>	sad_reflect	sad	reflect

Inspection of Table 5 reveals that  $V = 839$  for exact matches, which are subject to variation with respect to the tense of *be*. As expected, the  $\mathcal{P}$  score for exact matches is the highest in the list. Potential productivity is relatively weak with respect to Adj- $V_{\text{INF}}$  concatenations, i.e. once elements of variation have been removed. The same measure of productivity indicates that the Adj and  $V_{\text{INF}}$  slots are far less productive when they are considered individually than when they are combined. This is due to the verbatim recurrence of most verbs in the infinitive and, to a much lower extent, adjectives. Interestingly, global productivity scores are negatively correlated with  $\mathcal{P}$  scores. This is because  $\mathcal{P}$  it focuses exclusively on the probability of encountering new types. Not only does  $P^*$  account for the probability of observing new types, but it also includes observed types. The high  $P^*$  scores reflect the high influence of  $V$ .

Table 5: Productivity measures for exact matches, concatenated slots, the adjective slot, and the verb slot

	$N(C)$	$V$	$V1$	$\mathcal{P}$	$P^*$
exact matches	2136	839	639	0.2991573	2804.545
Adj- $V_{\text{INF}}$	2136	578	379	0.1774345	3257.541
Adj slot	2136	162	71	0.0332397	4873.69
$V_{\text{INF}}$ slot	2136	104	28	0.01310861	7933.714

Table 6 shows which LNRE model provides the best fit for each distribution. Each time, goodness-of-fit is evaluated with a multivariate  $\chi^2$  test. The fit is good when the  $\chi^2$  value is low and the corresponding  $p$ -value large (Baayen 2008, 233). Satisfactory models provide reliable interpolated and extrapolated expected values for the vocabulary size and the spectrum elements. These values are plotted in the form of VGCs in Figure 4.

Table 6: LNRE models and goodness-of-fit for exact matches, concatenated slots, the adjective slot, and the verb slot

	selected LNRE model	$\chi^2$	df	$p$
exact matches	finite Zipf-Mandelbrot	2.608132	4	0.6253833
Adj- $V_{\text{INF}}$	Generalized Inverse Gauss-Poisson	1.405234	5	0.9237399
Adj slot	finite Zipf-Mandelbrot	7.2188	3	0.06524139
$V_{\text{INF}}$ slot	Generalized Inverse Gauss-Poisson	3.762441	3	0.2882852

Figure 4 confirms what Table 5 has outlined, namely that the curves for exact matches and concatenated slots are steeper than the curves of Adj and  $V_{\text{INF}}$ . In fact, if we had three times the number of occurrences in a larger corpus containing data of the same type, productivity indexed on type and hapax legomena would still be expanding, as the steepness of the curves show. The *it* BE ADJ *to*  $V_{\text{INF}}$  *that* construction generates enough hapax legomena to be productive despite the limited number of occurrences (2136 tokens). All in all, multislot constructions can be productive even when their lexical constituents are not.

Something that the VGCs do not show is whether the lexical choice is mutually constrained between the Adj slot and the  $V_{\text{INF}}$  slot. This interdependence can be demonstrated as follows (Zeldes 2012, 130). Let  $P(HL_{\text{ADJ}})$  and  $P(HL_{\text{VINF}})$  denote the probabilities of finding hapax legomena in the Adj slot and  $V_{\text{INF}}$  slot respectively. If the probabilities are independent, we should expect  $P(HL_{\text{ADJ}} \cup HL_{\text{VINF}})$ , i.e. the probability of hapax legomena for Adj- $V_{\text{INF}}$ , to be one minus the product of the complementaries to  $\mathcal{P}$  in each slot. This is summarized in equation (viii):

$$P(HL_{\text{ADJ}} \cup HL_{\text{VINF}}) \approx 1 - (1 - \mathcal{P}_{\text{ADJ}}) \cdot (1 - \mathcal{P}_{\text{VINF}}) \quad (\text{viii})$$

We observe a 52.94% deviation between  $P(HL_{\text{ADJ}} \cup HL_{\text{VINF}})$  (0.18) and  $1 - (1 - \mathcal{P}_{\text{ADJ}}) \cdot (1 - \mathcal{P}_{\text{VINF}})$  (0.34). The lexical choice is mutually constrained in the two slots of the *it* BE ADJ *to*  $V_{\text{INF}}$  *that* construction. This implies that when we see a given slot (Adj or  $V_{\text{INF}}$ ), we have a reasonably strong expectation about the other slot.

What is left to see is the direction of this expectation. Because of the low productivity of  $V_{\text{INF}}$ , we might be inclined to say that given an adjective, the verb is easily predictable. However, we should not underestimate the high predictability of the adjective given a verb in some

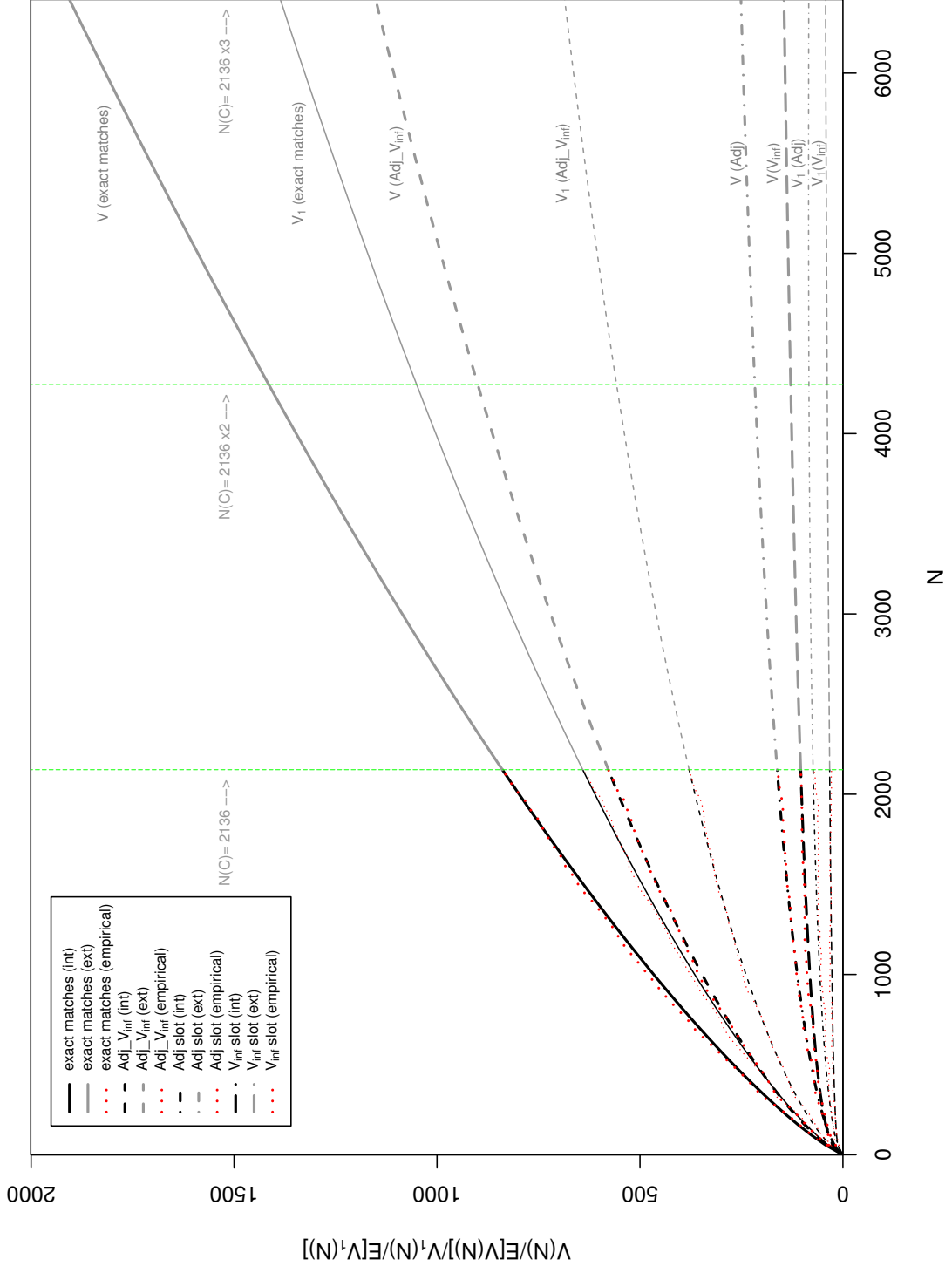


Figure 4: Vocabulary growth curves for exact matches, concatenated slots, the adjective slot, and the verb slot with interpolations (int) and extrapolations (ext) to three times the sample size

subschemas of the construction because not all adjectives are equally productive. To measure the asymmetric dependency at work in *it* BE ADJ *to* V<sub>INF</sub> *that*,  $\Delta P_{V_{INF}|ADJ}$  and  $\Delta P_{ADJ|V_{INF}}$  were computed for each A–V<sub>INF</sub> type. To better assess the detail of the asymmetries at work in Adj–V<sub>INF</sub> pairs, the pairwise differences of  $\Delta P$  values were calculated and compared to a well-known symmetric association measure:  $\chi^2$ . Table 7 summarizes the results.

Table 7: A comparison of symmetric and asymmetric association measures

a. Top 10 covarying collexemes				c. Top 10 negative $\Delta P$ values			
Adj	V <sub>INF</sub>	$\chi^2$	$\Delta P_{V_{INF} ADJ} - \Delta P_{ADJ V_{INF}}$	Adj	V <sub>INF</sub>	$\chi^2$	$\Delta P_{V_{INF} ADJ} - \Delta P_{ADJ V_{INF}}$
<i>misleading</i>	<i>claim</i>	141095.92	0.01	<i>necessary</i>	<i>confirm</i>	76.44	-0.96
<i>misleading</i>	<i>say</i>	19641.98	0.02	<i>necessary</i>	<i>recollect</i>	5363.83	-0.96
<i>difficult</i>	<i>believe</i>	16412.98	0.05	<i>tempting</i>	<i>fantasise</i>	25.14	-0.96
<i>important</i>	<i>remember</i>	10687.03	-0.38	<i>safe</i>	<i>wager</i>	0.46	-0.95
<i>straightforward</i>	<i>show</i>	10605.48	0.96	<i>sufficient</i>	<i>mention</i>	29.32	-0.95
<i>heartening</i>	<i>know</i>	8702.66	0.26	<i>essential</i>	<i>monitor</i>	1.30	-0.95
<i>necessary</i>	<i>recollect</i>	5363.83	-0.96	<i>possible</i>	<i>arrange</i>	70.33	-0.94
<i>unhistorical</i>	<i>assume</i>	4086.77	0.93	<i>possible</i>	<i>concede</i>	72.62	-0.94
<i>possible</i>	<i>specify</i>	3015.85	-0.93	<i>hard</i>	<i>disagree</i>	103.71	-0.94
<i>necessary</i>	<i>prove</i>	2859.31	-0.18	<i>reasonable</i>	<i>ask</i>	6.70	-0.94

b. Bottom 10 covarying collexemes				d. Top 10 positive $\Delta P$ values			
Adj	V <sub>INF</sub>	$\chi^2$	$\Delta P_{V_{INF} ADJ} - \Delta P_{ADJ V_{INF}}$	Adj	V <sub>INF</sub>	$\chi^2$	$\Delta P_{V_{INF} ADJ} - \Delta P_{ADJ V_{INF}}$
<i>correct</i>	<i>argue</i>	0.00	0.16	<i>unsound</i>	<i>argue</i>	25.36	0.96
<i>easy</i>	<i>tell</i>	0.00	-0.44	<i>dubious</i>	<i>argue</i>	2019.66	0.96
<i>hard</i>	<i>accept</i>	0.00	-0.14	<i>bizarre</i>	<i>suppose</i>	11.10	0.96
<i>heartening</i>	<i>find</i>	0.00	0.14	<i>arrogant</i>	<i>suppose</i>	11.77	0.96
<i>intriguing</i>	<i>hear</i>	0.00	-0.08	<i>rational</i>	<i>suppose</i>	1843.39	0.96
<i>possible</i>	<i>suppose</i>	0.00	-0.01	<i>insane</i>	<i>imagine</i>	0.88	0.96
<i>vital</i>	<i>remember</i>	0.00	0.12	<i>anachronistic</i>	<i>imagine</i>	0.99	0.96
<i>wrong</i>	<i>assume</i>	0.00	0.29	<i>ignorant</i>	<i>suggest</i>	0.06	0.96
<i>wrong</i>	<i>suppose</i>	0.00	0.03	<i>inadequate</i>	<i>suggest</i>	1.78	0.96
<i>difficult</i>	<i>convey</i>	0.01	-0.93	<i>shameful</i>	<i>suggest</i>	5.08	0.96

Tables 7a and 7b display the Adj–V<sub>INF</sub> pairs with the highest and the lowest  $\chi^2$  scores respectively. Although we might expect all pairs with extreme  $\chi^2$  values (high or low) to be characterized by low asymmetry due to their highly conventional nature and their fixedness, no such negative correlation is observed. Inspection of Tables 7c and 7d confirms the lack of correlation between  $\chi^2$  and  $\Delta P$ . In Table 7c, the verb is a much better predictor of the adjective than vice versa because the latter is found in more Adj–V<sub>INF</sub> pairs than the former and is therefore more productive, loosely speaking. Interestingly, the adjective types that are found in the list also happen to be among the most frequent first-slot constituents in the construction. Conversely, Adj is a much better predictor of V<sub>INF</sub> than vice versa in Table 7d. Four verb types are found in the list: *argue*, *imagine*, *suggest*, and *suppose*. These productive verbs are among the most frequent second-slot constituents in the construction.

Whether  $\Delta P$  has any bearing on the productivity assessment of *it* BE ADJ *to* V<sub>INF</sub> *that* remains to be shown. To summarize the contributions of all the abovementioned measures, the mean  $\Delta P$  and  $\chi^2$  scores for each adjective and verb found in the construction were collected, together with the corresponding mean  $\mathcal{P}$  and  $\mathcal{P}^*$  scores. The scores were gathered in a data frame, of which Table 8 is a snapshot. The data frame was submitted to principal component analysis (henceforth PCA). PCA consists in reducing the dimensionality of the input table by decomposing the total variance of the table into a limited number of components (for an overview, see Desagulier 2017, Section 10.2).

The 266 observations consist of all the adjective and verb types of *it* BE ADJ *to* V<sub>INF</sub> *that*. They were examined in the light of four quantitative variables ( $\Delta P_{DIFF}$ ,  $\chi^2$ ,  $\mathcal{P}^*$ , and  $\mathcal{P}$ ) and one illustrative variable (the category of the constituent). Because the variables are in different



Table 8: Productivity-related measures (snapshot)

word	category	$\Delta P$ diff	$\chi^2$	$\mathcal{P}^*$	$\mathcal{P}$
<i>able</i>	Adj	0.45	3.80	0.00	0.03
<i>absurd</i>	Adj	0.10	14.94	0.00	0.04
<i>accept</i>	V <sub>INF</sub>	0.15	35.05	0.00	0.25
<i>acceptable</i>	Adj	0.90	0.33	0.00	0.01
<i>accurate</i>	Adj	0.86	0.61	0.00	0.00
<i>acknowledge</i>	V <sub>INF</sub>	-0.22	13.08	0.00	0.07
<i>add</i>	V <sub>INF</sub>	-0.10	13.20	0.00	0.11
<i>adequate</i>	Adj	0.87	13.11	0.00	0.01
<i>admit</i>	V <sub>INF</sub>	0.50	3.80	0.00	0.07
<i>advisable</i>	Adj	0.44	7.26	0.00	0.03
...	...	...	...	...	...

units, they were centered and standardized. Negative  $\Delta P_{\text{DIFF}}$  scores were switched to absolute values to avoid spreading points on either side of a component while preserving the magnitude of the measure.

Inspection of the graph of variables in the left part of Figure 5, reveals that the first component (i.e. the horizontal axis) accounts for 54.03% of the variance. It is positively correlated with  $\mathcal{P}$  ( $cor=0.949$ ) and  $\mathcal{P}^*$  ( $cor=0.928$ ) and negatively correlated with  $\Delta P_{\text{DIFF}}$  ( $cor=-0.628$ ). This is useful for interpreting the graph to the right. The further right along the horizontal axis, the more productive the constructional constituents from a hapax-based viewpoint.

The second component (i.e. the vertical axis) accounts for 24.94% of the variance. It is positively correlated with  $\chi^2$  ( $cor=0.997$ ). The further up along the vertical axis, the stronger the symmetric association between the constituent (adjective or verb) and its counterpart.

The graph in the right-hand side of Figure 5 displays the constructional constituents in the two-dimensional space spanned by the first two components. So as not to clutter the graph, only the top 25% of words that contribute the most to each component were projected. The other observations appear as dots without a label.

The most productive constituents from a hapax-based perspective appear neatly in the right part of the plane representation of individuals in Fig. 5. They belong to weakly associated pairs. Tab. 9 displays the 20 construction constituents that contribute the most to the first component (Dim 1, horizontal axis). They are the 20 most productive words based on hapax-based measures. The words are sorted in descending order according to their position on the axis (coord) and their contribution to the component (contrib).<sup>3</sup> The last column features the quality of representation of the word to the component ( $\cos^2$ ).<sup>4</sup> The most productive constituents are mostly Perception–Cognition–Utterance verbs (Givón 2001). The five adjectives in the list denote easiness/difficulty (*easy*, *difficult*, *hard*), or epistemic or deontic meaning (*possible*, *essential*). Verbs are more likely than adjectives to occur in nonce subschemas in the *it* BE ADJ to V<sub>INF</sub> *that* construction.

$\Delta P_{\text{DIFF}}$  is well represented in the third component only ( $cor=0.778$ ). The right plot in Figure 6 displays the constructional constituents in the two-dimensional space spanned by the second and third components. So as not to clutter the graph, only the 10 observations that have the

<sup>3</sup>The contribution of a construction constituent to a component is a measure of how much the constituent affects the construction of the component.

<sup>4</sup>The quality of representation of a construction constituent on a component is measured via the percentage of inertia of the constituent projected on the component.

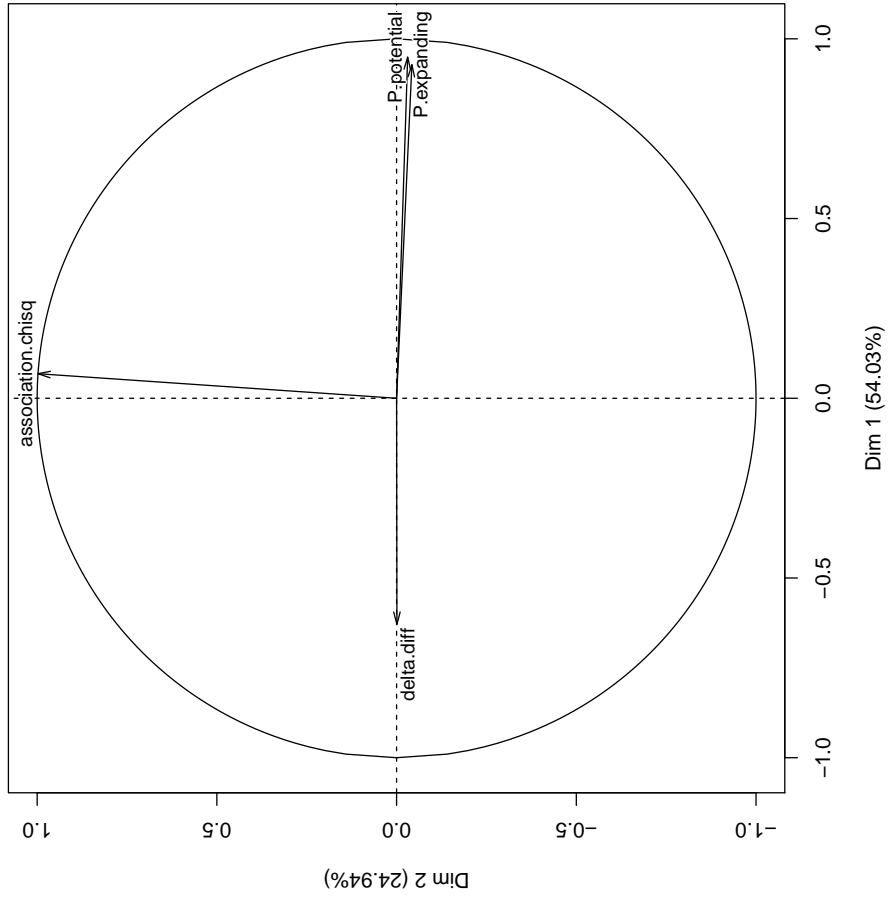
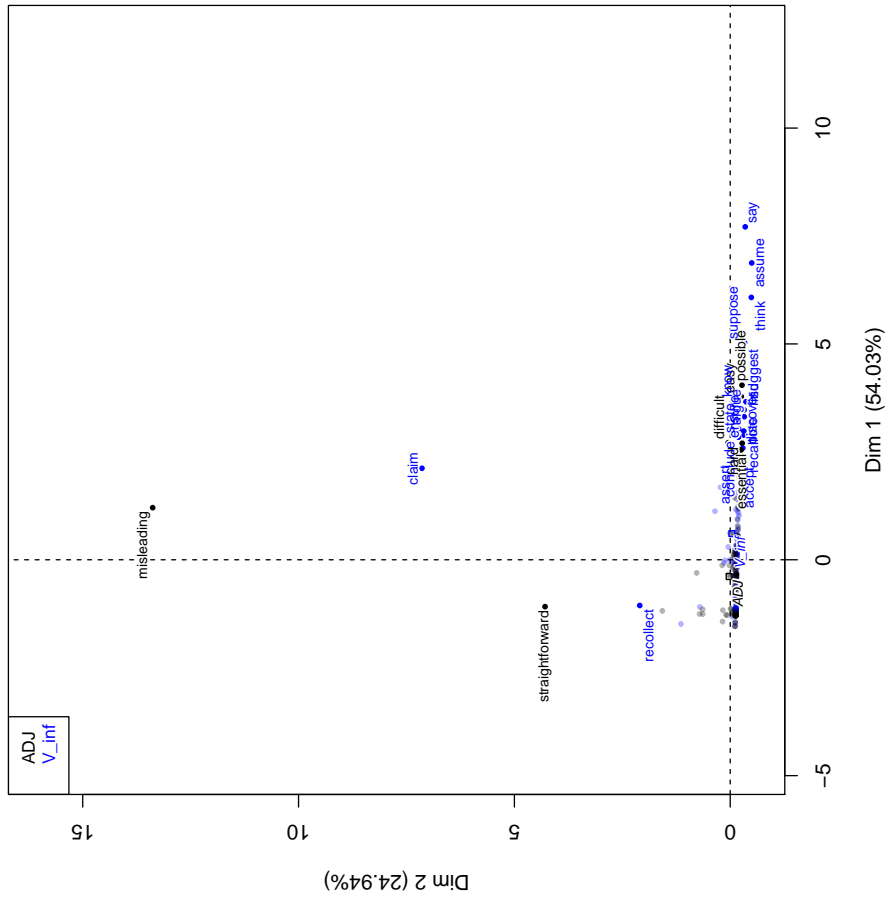


Figure 5: PCA output (dimensions 1 & 2) – left: active variables;  
right: plane representation of individuals

Table 9: The 20 most productive constituents according to hapax-based measures

word	category	coord	contrib	cos <sup>2</sup>
<i>say</i>	V <sub>INF</sub>	7.712	10.35	0.838
<i>assume</i>	V <sub>INF</sub>	6.875	8.22	0.836
<i>think</i>	V <sub>INF</sub>	6.077	6.42	0.842
<i>suppose</i>	V <sub>INF</sub>	5.027	4.4	0.897
<i>possible</i>	Adj	4.045	2.85	0.781
<i>easy</i>	Adj	3.783	2.49	0.788
<i>state</i>	V <sub>INF</sub>	3.723	2.41	0.893
<i>know</i>	V <sub>INF</sub>	3.669	2.34	0.874
<i>suggest</i>	V <sub>INF</sub>	3.66	2.33	0.868
<i>find</i>	V <sub>INF</sub>	3.473	2.1	0.79
<i>ensure</i>	V <sub>INF</sub>	3.314	1.91	0.898
<i>argue</i>	V <sub>INF</sub>	3.027	1.59	0.948
<i>note</i>	V <sub>INF</sub>	2.98	1.54	0.931
<i>recall</i>	V <sub>INF</sub>	2.887	1.45	0.898
<i>conclude</i>	V <sub>INF</sub>	2.841	1.4	0.977
<i>difficult</i>	Adj	2.712	1.28	0.813
<i>hard</i>	Adj	2.699	1.27	0.805
<i>discover</i>	V <sub>INF</sub>	2.588	1.17	0.945
<i>essential</i>	Adj	2.545	1.13	0.827
<i>assert</i>	V <sub>INF</sub>	2.287	0.91	0.976

highest contribution on the two dimensions of the plot have a label. The third component (Dim 3, vertical axis) accounts for 18.59% of the variance. It is also partially correlated with  $\mathcal{P}$  ( $cor=0.301$ ) and  $\mathcal{P}^*$  ( $cor=0.219$ ).

Table 10 displays the 20 construction constituents that contribute the most to the third component. They correspond to words that are characterized by the highest  $\Delta P_{DIFF}$  scores. These words cluster in the upper part of the plot, along the vertical axis.

The subschemas that are the most asymmetric involve verbs mostly. These verbs can be said to induce productivity because they combine with a wide variety of adjectives. Ten constituents in Tab. 10 are also found in Tab. 9: *say*, *assume*, *think*, *find*, *suppose*, *know*, *suggest*, *state*, *ensure*, and *possible*. This means that productive constituents combine with a wider variety of adjectives or verbs, and that sometimes these combinations are found only once. Inspection of the third component allows us to conclude that hapax-based measures and  $\Delta P$  highlight different yet related aspects of productivity.

## 6 Discussion

The above suggests that sticking to the schematic level is not enough to understand how *it BE ADJ to V<sub>INF</sub> that* works. The VGCs in Fig. 4 show that the construction is productive at the schematic levels (exact matches and Adj-V<sub>INF</sub>). Although the productivity scores are modest (e.g.  $\mathcal{P}_{exact\ matches} \approx 0.3$  and  $\mathcal{P}_{Adj-V_{INF}} \approx 0.18$ , see Tab. 5), the corresponding curves are steep. Their shapes suggest that productivity has not reached its peak and that we would observe more and more new types if we were to inspect more corpus data of the same kind.

In contrast, the flat curves indicate that the hapax-based productivity is far from obvious

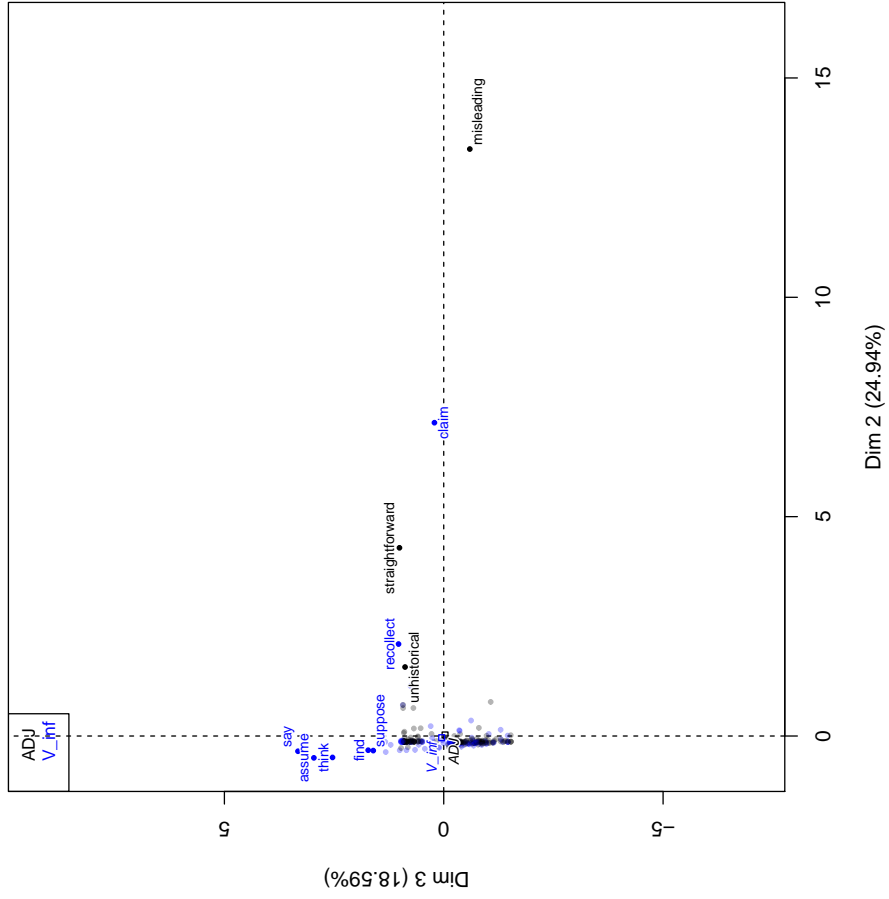
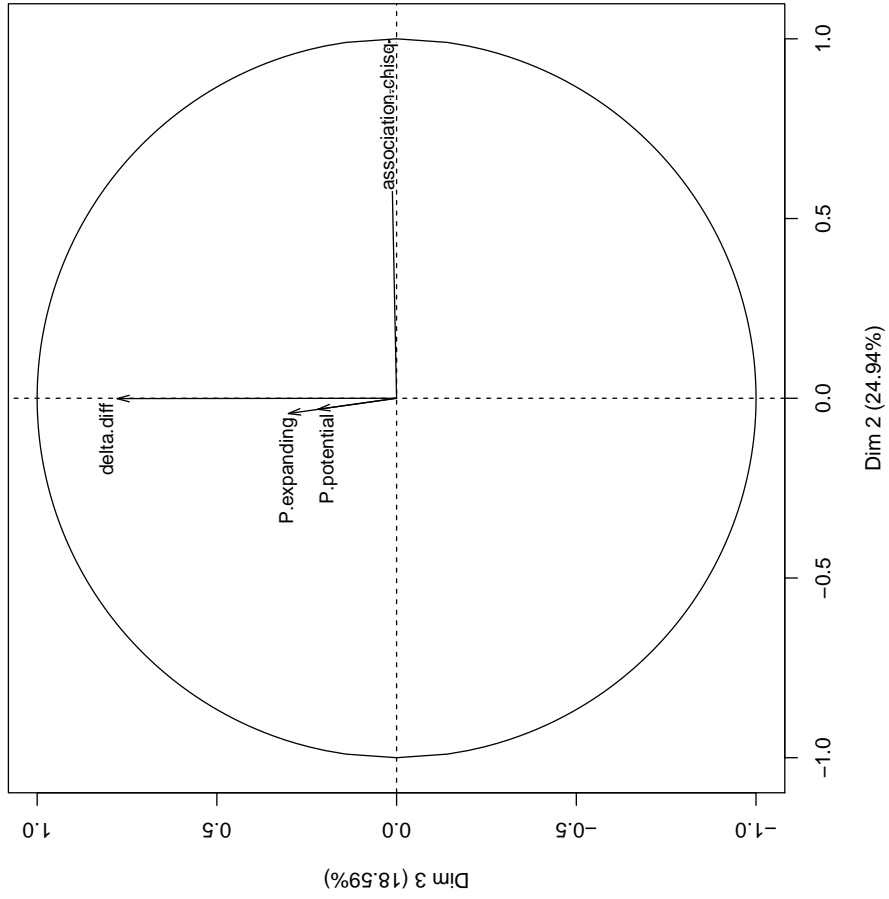


Figure 6: PCA output (dimensions 2 & 3) – left: active variables;  
right: plane representation of individuals

Table 10: The 20 most asymmetric constituents according to  $\Delta P$  difference

word	category	coord	contrib	$\cos^2$
<i>say</i>	V <sub>INF</sub>	3.323	5.58	0.156
<i>assume</i>	V <sub>INF</sub>	2.96	4.43	0.155
<i>think</i>	V <sub>INF</sub>	2.536	3.25	0.147
<i>find</i>	V <sub>INF</sub>	1.723	1.5	0.195
<i>suppose</i>	V <sub>INF</sub>	1.607	1.3	0.092
<i>know</i>	V <sub>INF</sub>	1.332	0.9	0.115
<i>suggest</i>	V <sub>INF</sub>	1.323	0.88	0.113
<i>state</i>	V <sub>INF</sub>	1.21	0.74	0.094
<i>recollect</i>	V <sub>INF</sub>	1.031	0.54	0.161
<i>straightforward</i>	Adj	1.01	0.52	0.05
<i>ensure</i>	V <sub>INF</sub>	1.004	0.51	0.082
<i>confirm</i>	V <sub>INF</sub>	0.979	0.48	0.412
<i>fantasise</i>	V <sub>INF</sub>	0.977	0.48	0.41
<i>possible</i>	Adj	0.972	0.48	0.045
<i>wager</i>	V <sub>INF</sub>	0.972	0.48	0.408
<i>mention</i>	V <sub>INF</sub>	0.971	0.48	0.409
<i>monitor</i>	V <sub>INF</sub>	0.961	0.47	0.405
<i>arrange</i>	V <sub>INF</sub>	0.948	0.45	0.402
<i>concede</i>	V <sub>INF</sub>	0.948	0.45	0.403
<i>disagree</i>	V <sub>INF</sub>	0.941	0.45	0.401

at the level of the adjective or the verb slots (regardless of the other constituents that they pair with). It means that the pool of adjectival and verbal constituents that speakers tap into to assemble new Adj-V<sub>INF</sub> types is relatively limited, which might be due to the small sample size. However limited, the pool of constituents generates a handful of subschemas that are undeniably productive. Each productive subschema is indexed on either (a) an adjective denoting easiness/difficulty (*easy*, *difficult*, *hard*) or modal meaning (*possible*, *essential*), or (b) a Perception-Cognition-Utterance verb (*say*, *assume*, *think*, *suppose*, etc., see Tab. 9).

$\Delta P$  is not meant to replace more conventional productivity measures. As we have seen, no direct correlation was found between  $\Delta P$  scores and the scores based on hapax-based measures. Whether a slot is actually productive cannot be concluded from an inspection of  $\Delta P$  scores only. However, what  $\Delta P$  highlights is whether a particular slot is a potential locus of productivity (see Tabs. 7c and 7d).

Not all Adj-V<sub>INF</sub> pairings are attested. Selectional restrictions apply. One such restriction is the semantic affinity between the adjective and the verb. For example, knowing something is hardly ever seen as stupid. Therefore, the token *it is stupid to know* that is most unlikely to occur. Selectional preferences also apply. Given that the construction is used for stance marking in formal, written contexts (see Fig. 1), usage bias may reflect in speakers' preferences for some adjectives and some verbs. I assume that this combination of selectional restrictions and preferences shapes the internal architecture of the *it BE ADJ to V<sub>INF</sub> that* construction. Pending experimental validation, speakers are sensitive to the internal composition of the construction in the sense that they either use fixed Adj-V<sub>INF</sub> combinations or assemble constituents on the spot to fill their communication needs.

## 7 Conclusion

The paper has examined distributional characteristics of the *it* BE ADJ to  $V_{\text{INF}}$  *that* pattern, as attested in the BNC. The theoretical goal of the paper was to assess existing CxG accounts for the adequacy of their definition of construction. I have employed a combination of metrics (association measures and hapax-based measures) to evaluate productivity of the pattern. The results show that the pattern in question is partly productive, especially with specific adjective and verbs.

Whether *it* BE ADJ to  $V_{\text{INF}}$  *that* is a pattern or a construction is, after all, a matter of theoretical preferences, as I have already argued elsewhere (Desagulier 2015). Because I endorse the tenets of CCG, the constructional status of *it* BE ADJ to  $V_{\text{INF}}$  *that* is unproblematic from the start because productivity is not key in determining whether a pattern is a construction or not. However, if one is willing to maintain the distinction between a pattern of coining and a construction, one should at least not relegate patterns of coining to an ancillary status on the grounds that their contribution to our understanding of linguistic competence and performance is null.

Without endorsing the tenets of non-redundant CxG (especially productivity as a criterion for constructionality), I borrowed them, operationalized productivity in the framework of established measures, and showed that different conclusions can be made. Even if one were to admit that patterns of coining are at the periphery of the mental inventory of constructions (also known as ‘the constructicon’), they do serve as templates for the formation of new types.

## References

- Aijmer, K. (2009). Seem and evidentiality. *Functions of Language*, 16(1), 63–88.
- Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, 15(3), 147–149.
- Baayen, R. H. (1989). *A corpus-based approach to morphological productivity. Statistical analysis and psycholinguistic interpretation*. Amsterdam: Centrum Wiskunde en Informatica.
- Baayen, R. H. (1993). On frequency, transparency and productivity. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1992* (pp. 181–208). Dordrecht & London: Kluwer.
- Baayen, R. H. (2001). *Word frequency distributions*. Dordrecht: Kluwer Academic Publishers.
- Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.
- Baayen, R. H. (2009). Corpus linguistics in morphology: Morphological productivity. In A. Lüdeling & M. Kytö (Eds.), *Corpus linguistics. An international handbook* (Vol. 2, pp. 899–919). Berlin: Mouton de Gruyter.
- Baayen, R. H. & Lieber, R. (1991). Productivity and English derivation: A corpus-based study. *Linguistics*, 29, 801–843.
- Baratta, A. M. (2009). Revealing stance through passive voice. *Journal of Pragmatics*, 41(7), 1406–1421.
- Biber, D. (2006). Stance in spoken and written university registers. *Journal of English for Academic Purposes*, 5(2), 97–116.
- Biber, D. & Finegan, E. (1988). Adverbial stance types in English. *Discourse Processes*, 11(1), 1–34.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E., & Quirk, R. (1999). *Longman Grammar of Spoken and Written English*. MIT Press Cambridge, MA.
- Biq, Y.-O. (2004). Construction, reanalysis, and stance: ‘*v yi ge n*’ and variations in mandarin chinese. *Journal of Pragmatics*, 36(9), 1655–1672.

- Boas, H. (2003). *A constructional approach to resultatives*. Stanford: CSLI Publications.
- Bybee, J. (2010). *Language, usage, and cognition*. Cambridge: Cambridge University Press.
- Chang, L.-H. (2009). Stance uses of the mandarin LE constructions in conversational discourse. *Journal of Pragmatics*, 41(11), 2240–2256.
- Charles, M. (2007). Argument or evidence? disciplinary variation in the use of the noun that pattern in stance construction. *English for Specific Purposes*, 26(2), 203–218.
- Dancygier, B. (2012). Negation, stance verbs, and intersubjectivity. *Viewpoint in language: A multimodal perspective*, 69.
- Dancygier, B. & Sweetser, E. (2012). *Viewpoint in language: A multimodal perspective*. Cambridge University Press.
- Desagulier, G. (2015). A lesson from associative learning: Asymmetry and productivity in multiple-slot constructions. *Corpus Linguistics and Linguistic Theory*. doi:10.1515/cllt-2015-0012
- Desagulier, G. (2017). *Corpus linguistics and statistics with R. Introduction to quantitative methods in linguistics*. Quantitative Methods in the Humanities and Social Sciences. New York: Springer.
- Diani, G. (2008). Emphasizers in spoken and written academic discourse: The case of really. *International Journal of Corpus Linguistics*, 13(3), 296–321.
- Ellis, N. (2006). Language acquisition as rational contingency learning. *Applied Linguistics*, 27(1), 1–24.
- Ellis, N. & Ferreira-Junior, F. (2009). Constructions and their acquisition: Islands and the distinctiveness of their occupancy. *Annual Review of Cognitive Linguistics*, 7, 187–220.
- Englebretson, R. (2007). *Stancetaking in discourse: Subjectivity, evaluation, interaction*. Amsterdam ; Philadelphia: John Benjamins Publishing.
- Fillmore, C. (1997). *Construction Grammar lecture notes*. Retrieved August 14, 2015, from <http://www.icsi.berkeley.edu/~kay/bcg/lec02.html>
- Fillmore, C., Kay, P., & O'Connor, C. (1988). Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language*, 64(3), 501–538.
- Givón, T. (2001). *Syntax: An introduction*. John Benjamins Publishing.
- Goldberg, A. E. (1995). *Constructions: A construction grammar approach to argument structure*. Cognitive theory of language and culture. Chicago: University of Chicago Press.
- Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences*, 7(5), 219–224.
- Goldberg, A. E. (2006). *Constructions at work: The nature of generalization in language*. Oxford & New York: Oxford University Press.
- Goldberg, A. E. (2009). The nature of generalization in language. *Cognitive Linguistics*, 20(1), 93–127.
- Gray, B. & Biber, D. (2014). Stance markers. In K. Aijmer & C. Rühlemann (Eds.), *Corpus pragmatics: A handbook* (pp. 219–248). Cambridge University Press.
- Gries, S. T. (2013). 50-something years of work on collocations: What is or should be next... *International Journal of Corpus Linguistics*, 18(1), 137–166.
- Gries, S. T. & Stefanowitsch, A. (2004). Co-varying collexemes in the *into*-causative. In M. Achard & S. Kemmer (Eds.), *Language, culture, and mind* (pp. 225–236). Stanford: CSLI.
- Hewings, M. & Hewings, A. (2002). “It is interesting to note that...”: A comparative study of anticipatory ‘it’ in student and published writing. *English for Specific Purposes*, 21(4), 367–383.
- Hunston, S. (2007). Using a corpus to investigate stance quantitatively and qualitatively. In R. Englebretson (Ed.), *Stancetaking in discourse: Subjectivity, evalu-*

- ation, interaction (Chap. 2, pp. 27–48). Amsterdam ; Philadelphia: John Benjamins Publishing.
- Hunston, S. (2011). *Corpus approaches to evaluation: Phraseology and evaluative language*. London: Routledge.
- Hyland, K. & Tse, P. (2005). Hooking the reader: A corpus study of evaluative that in abstracts. *English for specific purposes*, 24(2), 123–139.
- Kay, P. (2013). The limits of (Construction) Grammar. In T. Hoffmann & G. Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 32–48). Oxford: Oxford University Press.
- Kay, P. & Fillmore, C. (1999). Grammatical constructions and linguistic generalizations: The *What's X doing Y?* construction. *Language*, 75, 1–33.
- Martin, J. R. & White, P. R. (2005). *The language of evaluation*. Basingstoke: Palgrave Macmillan.
- Morita, E. (2015). Japanese interactional particles as a resource for stance building. *Journal of Pragmatics*, 83, 91–103.
- Myers, G. & Lampropoulou, S. (2012). Impersonal you and stance-taking in social research interviews. *Journal of Pragmatics*, 44(10), 1206–1218.
- Rescorla, R. A. (1968). Probability of shock in the presence and absence of CS in fear conditioning. *Journal of Comparative and Physiological Psychology*, 66, 1–5.
- Stefanowitsch, A. & Gries, S. T. (2005). Co-varying collexemes. *Corpus Linguistics and Linguistic Theory*, 1(1), 1–46.
- Wagner, A. R. & Rescorla, R. A. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II* (pp. 64–99). New York: Appleton-Century-Crofts.
- Zeldes, A. (2012). *Productivity in argument selection: From morphology to syntax*. Berlin & New York: Mouton de Gruyter.