

Entrenchment inhibition: Constructional change and repetitive behaviour can be in competition with large-scale ‘re-compositional’ creativity

Abstract

This paper addresses creativity as inhibition of repetitive behaviour. We argue that entrenchment and constructional change (e.g. Traugott and Trousdale 2013) can be in competition with large-scale creative attempts of re-composition of constructions’ internal constituency. After undergoing chunking, the recurrent usage of a construction may be significantly counter-balanced with new attempts of entrenchment inhibition (viz. inhibition of entrenchment) (Tantucci et al. 2018). These are cases where speakers opt for more compositional and less predictable ways to express a similar meaning of a conventionalised form. We focus on the constructionalization of Noun-participle compounds (e.g. *snow-covered*) in the Historical Corpus of American English (COHA, Davis 2012). During the second part of the 20th century, speakers increasingly inhibit the usage of conventionalised NP-Pp forms in favour of more compositional strategies involving the same internal constituents. This entails that constructional change not only affects the meaning of the chunk that undergoes constructionalization, but also the way speakers creatively re-discover its internal constituency. These results additionally aim to inform research in cognitive architectures and artificial intelligence, where creativity is often merely considered as a problem-solving mechanism rather than a potential process of inhibition of automatised behaviour.

1 Introduction

Language change contributes to the theoretical understanding of natural linguistic systems, but also crucially informs research in different spheres of cognitive science and applied linguistics (i.e. Ellis and Larsen 2006; Diessel 2011). This study suggests that **entrenchment** (e.g. Langacker 1987, Schmid 2017) and chunking (cf. Newell 1990, Bybee 2010; Ellis 2017) of an item undergoing constructionalization are not necessarily incremental. After reaching the highest degree of entrenchment, a construction may start to be in competition with large-scale mechanisms of semantic re-composition of its internal constituency, viz. speakers creatively opting for an ad-hoc greeting such as *I wish you an truly good morning*, instead of the more entrenched *good morning*. We define this creative process as **entrenchment inhibition** (cf. Tantucci et al. 2018) and provide a corpus-based case study from the COHA (Corpus of Historical English, cf. Davis 2012) centred on

the constructionalization of Noun phrase-Past participle compounds (henceforth NP-Pps, e.g. [*snow-covered*]) in American English from 1810 up to the present. Namely, entrenchment inhibition involves individual creative attempts to re-compose meanings and forms of already entrenched constructions. Entrenchment inhibition normally does not lead to innovation. It is rather part of the ecosystem of a conventionalised construction, as it occurs as large-scale, ad-hoc forms of alteration that inherently balance the repetitive behaviour of entrenched linguistic forms. This project recalls influential early work suggesting that a desire for emphasis leads to additional elements that combine compositionally, e.g. the addition of *pas and rien to ne* in French (Trudgill 1992)

We discuss data showing that the entrenchment inhibition of NP-Pp constructs such as *Washington-based* or *tree-lined* occurs significantly at a stage in time when the NP-Pp node has already reached the highest degree of schematicity. At this point, entrenchment inhibition comes into play as a competing mechanism counterbalancing the automatization and repetition of the NP-Pp forms. From a corpus-based analysis of 8781 annotated occurrences will emerge that, after the entrenchment and constructional formation of NP-Pps, people will start to combine the internal constituents of the construction with novel strategies (*shaped in the form of a U* - COHA/1946, in the place of [*U-shaped*]) that are not found in the first century of the COHA. Importantly, there cannot be a perfect equivalence between the meaning of a chunked form and a corresponding de-entrenched usage. For instance, the NP-Pp compounding can be coordinated with adjectives as in *efficient* and *market-oriented approaches*. From a radical construction grammar angle, here actions intersect with the propositional act of modification (cf. Croft 2001: 66). The same is not true for periphrastic usages of the same constituents, e.g. *efficient approaches that were oriented to the market*. While a speaker may have opted for the former chunked expression ([*market-oriented*]), s/he may otherwise refer to the same constituents expressing a similar, yet not perfectly equivalent meaning (as in the case of *efficient approaches that were oriented to the market*). We argue that this is indeed one of the reasons that prompt entrenchment inhibition, viz. speakers/writers' intention to express a surplus of meaning, despite the presence of an entrenched form that would require less executive functioning. Mismatches in information structure, compositionality and syntactic patterning are arguably thus what motivates speakers/writers to potentially opt for longer and more complex forms including the same constituents that can be found in more entrenched construct(ion)s.

The corpus-based results of this paper crucially suggest that linguistic creativity is not only an exceptional mechanism of extravagant innovation (e.g. Haspelmath 1999), triggering subsequent reanalysis in the linguistic system as a whole. Crucially, creativity is also a very common process of

mitigation of already conventionalised and repetitive behaviour. This is a point that seems to have been partly overlooked in the usage-based literature and in research of cognitive architectures and artificial intelligence.

This paper is structured as follows: section 2 introduces entrenchment, chunking and constructionalization as they have been discussed in the usage-based literature. Sections 2.1 and 2.2 problematise the notion of on-line creativity in usage-based linguistics and cognitive architecture research. In the same sections we introduce the notion of entrenchment inhibition with the aim of addressing creativity not only as a problem-solving process of innovation, but also as a large-scale mechanism of inhibition of automatised behaviour. Section 3 introduces the Noun-participle compounds (henceforth NP-Pps) and posits the two main research questions of this study:

1. whether entrenchment inhibition can significantly act as competing force of entrenchment through a process of constructionalization and whether this can be proved statistically,
2. whether entrenchment inhibition creatively intervenes on NP-Pps which did not collocate prior to their formation as compounds. The latter condition will be addressed in terms of absence of **pre-analysis** and will indicate that constructionalization itself has an effect on creativity, as completely unprecedented collocations will be possible after the formation of new NP-Pps (e.g. *oriented not so much to casual consumers* - COHA/2003 after the formation of [*consumer-oriented*]).

Section 4 provides the results of a distinctive collexeme analysis of the constructional change of NP-Pps throughout the COHA, including entrenchment inhibition as a competing variable. In particular, section 4.3 shows how a retrospective prediction of entrenchment merely based on isolated frequency of the NP-Pps compounds is possible, while a retrospective forecast of the competition between entrenchment and periphrastic strategies of entrenchment inhibition confirms a less linear and incremental and inferable trajectory of the constructional change of the NP-Pp compound. Section 5 digs further into the entrenchment inhibition of NP-Pps and unveils how entrenchment inhibition intersects NP-Pps without a pre-analytical history. Section 6 reports the conclusions of this study.

2 The role of entrenchment and chunking in the usage-based model

Most usage-based accounts centred on language change are primarily concerned with increasing tendencies towards the repetition and the predictability of verbal experience. Phenomena under enquiry often intersect with entrenchment (cf. Langacker 1987: 59; Croft 2000: 38; Zima and Brone 2015: 488), increase of schematicity (i.a. Bybee 2010; Traugott and Trousdale 2013: 22; Schmid 2017), chunking (Bybee 2010), and conventionalization (i.a. Traugott and Dasher 2002; Terkourafi 2015; Tantucci 2013, 2017a, 2017b) with a special emphasis given to the diachronic relationship between repetition and ‘bottom-up’ constructional abstraction (i.e. Hilpert 2015). In this study we also endorse the view of language as an adaptive system (cf. Beckner et al. 2009) that moves towards the uniformity, the automatised and predictability (cf. Bybee 2010) of verbal behaviour. Yet, we additionally aim at complementing the usage-based model by also taking into account large-scale creative inhibition of entrenched constructions.

Cognitive linguistics research has shown that function-specific chunks of verbalisation inherently affect how language is acquired, used and the way the linguistic system as a whole changes diachronically. Increased repetition of formulaic utterances leads to conventionalization (Bybee 1998; Heine and Kuteva, 2007; Tantucci 2015; Terkourafi 2015) of fixed patterns in individuals’ memory. Further changes then may occur formally at the phonetic, semantic, grammatical and especially pragmatic level: “speakers’ behaviour is based on their past interactions, and current and past interactions together feed forward into future behaviour” and “the structures of language emerge from interrelated patterns of experience, social interaction, and cognitive processes” (Beckner et al. 2009: 2). Usage-based research commonly assumes the probabilistic nature of linguistic behaviour and the emergence of chunked regularities from the interaction of agents in language use. Emphasis on the predictability of verbal experience underpins the probability of the word given the preceding or following word or words, and likelihood of the word based on the topic of the conversation (Gregory et al. 1999, Jurafsky et al. 2001, Jurafsky et al. 2002).

Token frequencies of linguistic constructions correlate with **entrenchment** (Croft 2000; Schmid 2007; Gries et al. 2010), which corresponds to the degrees of cognitive routinisation of linguistics structures and their likelihood to be stored in memory (cf. Langacker, 1987, 2009). In corpus linguistics, a similar tendency is addressed by the idiom principle: “a language user has available to him or her a large number of semi-preconstructed phrases that constitute single choices, even though they might appear to be analysable into segments” (Sinclair, 1991: 110). The same phenomenon is traditionally observed through the lenses of the so-called conventional symbolic units (Langacker, 1987; Croft and Cruse 2004), viz. structures “that a speaker has mastered quite

thoroughly, to the extent that he can employ it in largely automatic fashion, without having to focus his attention specifically on its individual parts for their arrangement [...] he has no need to reflect on how to put it together” (Langacker, 1987: 57).

Diachronic research has been increasingly addressing constructional and semantic change as a process of **chunking** (cf. Newell 1990; Bybee 2010). The latter is identified as “the underlying cognitive basis for morphosyntax and its hierarchical organization [...] of sequential experiences” which occurs mainly with repetition (Haiman 1994; Bybee, 2003, 2010: 34). Chunking leads to formation of formulaic or prefabricated sequences of words such as *take a break*, *break a habit*, *pick and choose* (Bybee 2002, 2010), and automatised processing progressively allowing co-articulation and reduction, as in the constructs *I don't know/I'm going* to grammaticalizing into more entrenched constructions *I dunno/I'm gonna*. Chunking arises with entrenchment and leads to progressive diminishing of the internal constituency and the compositionality of frequently used constructs. Such process underpins constructionalization when a new form-meaning pairing has been developed and widely recognised within a community of speakers (i.e. Traugott and Trousdale 2013).

2.1 Creativity as innovation towards predictable behaviour

How does the usage-based model account for creativity? In the usage-based literature, creativity comes into play as data-driven relationship between **innovation** and **change** (cf. Traugott and Trousdale 2013: 2). In fact, innovation becomes relevant when it is “replicated across populations of speakers resulting in conventionalisation” (Ibid; see also Weinreich et al. 1968; Andersen 2001). Creativity thus becomes important when replication of a new variant shifts from first to second and finally third order variation (cf. Croft 2000, 2010), or in other words, when a new form that is created by an individual progressively spreads through a population as a whole.

This paradigm reflects the emphasis that research in cognitive neuroscience has been placing on ‘creativity’ as an experience-based mechanism of computation geared towards learning/solving tasks and improving future online performance and predictability (cf. McRae et al. 1997; Roland et al. 2012 on predictions about upcoming linguistic material). Based on the same assumption, cognitive architectures and artificial intelligence models have been distinctively centred on experiential learning processes and subsequent memory retrieval. ACT-R (Anderson et al. 2004), SOAR systems (e.g. Newell 1990) or Icarus (Langley et al. 2004) are all cognitive architectures that

foreground problem-solving processing depending on previous chunks of experience. They all implement a responsive performance during an online task as a result of previously entrenched structures of behaviour (i.e. underpinning implicit memory) or entrenched propositional beliefs (i.e. having to do with declarative memory).

Crucially, this computational paradigm never includes online decisions of **inhibiting** patterns of experience otherwise successful in accomplishing tasks or solving problems. However, inhibitory control is a major area of research in cognitive science as well as neuroscience, and creativity is often thought to require inhibition of habits (Wood and Neal 2007; Trude and Nozari 2017). In fact, inhibition of entrenched patterns of verbal experience is precisely what humans often do during speech events, even when the contextual or preparatory conditions (cf. Searle, 1969) remain constant. For example, aware of an entrenched chunk *x* [*see you later*], Sp/Wr¹ may otherwise decide to utter *x*+/-*y* [*I'll see you again young man*] (BNC G5E PS285) despite *x* having been repeatedly proved to be felicitous in the same contextual conditions (cf. Tantucci et al. 2018)².

2.2 Creativity as inhibition of predictable behaviour: Entrenchment inhibition

Speakers' mastery of novel utterances is indeed an important issue in the linguistics literature (e.g. Lakoff 1970; Brown and Hanlon 1970; Braine 1971; Baker 1979; Bowerman 1988; Pinker 1989; Goldberg 1995; Gennari and Macdonald 2008). In some cases, this is argued to be primarily due to over-generalisation decreasing as one encounters more and more grammatical uses. (Braine and Brooks 1995; Ambridge et al. 2008). Nonetheless Harmon and Kapatsinski (2017) provide compelling evidence suggesting that frequent forms are preferentially extended to novel related uses. Through a number of exposure trials they show that accessibility results in use of frequent forms to express meanings that were related but not identical to the specific meanings paired with the forms in the participants' experience.

In other accounts, the notion of statistical preemption is proposed, suggesting that people learn novel constructs such as argument structure restrictions through indirect negative evidence (Boyd and Goldberg 2011; Clark 1987; Di Sciullo and Williams 1987; Foraker et al. 2007; Goldberg 1993, 1995, 2006, 2011; Pinker 1981). Stefanowitsch (2008) proposes a collostructional

¹ Speaker/Writer.

² From the demographically sampled section of the BNC, *see you later* appears in 29.9 instances per million words in the section starting from 45 year old up to 60+. It can thus be considered as a frequent expression among older generations, which allows to control the competition between *I'll see you again young man* and *I'll see you later* among speaker of the same age-range.

method accounting for the same phenomenon, namely what he calls negative entrenchment. Statistical preemption also places predictability at the core of its enquiry, suggesting that speakers attempt to anticipate others' utterances as they experience them (i.a. Allopenna et al. 1998; Kamide et al. 2003; Gibson et al. 2013; Jaeger and Snider 2013). Robenalt and Goldberg (2015) make the case for statistical preemption inducing speakers to prefer what is frequent, opting for familiar formulations to novel ones and making creative choices only in absence of available alternatives.

We similarly agree that speakers do statistically favour frequency and entrenchment over choices that may be cognitively more demanding. However, 'chunked' behaviour may not necessarily lead to continuous large-scale reiteration of the same entrenched item during verbal experience. It has been argued that episodic memory for words or lexical associations can be impaired by the previous retrieval of a related memory (e.g. Anderson, Bjork, and Bjork, 1994; but see also Anderson and Neely, 1996). Alternatively, it is proposed that such inhibition results from suppressing previous competitors (also defined as unlearning; e.g. Norman, Newman, and Detre 2007). In an influential account, Oppenheim et al. (2010) emphasise the role of experience in inhibition phenomena during language processing.

This paper aims at accounting for the competition (if any) between a speaker's inclination to favour a less entrenched strategy (e.g. *I'll see you again young man*) and the otherwise more predictable and conventional form available (e.g. *see you later*). The former occurring on a large-scale is what we define as **entrenchment inhibition**, which we illustrate in a case study centred on noun-participle compounding in American English. This survey shows that after the NP-Pp construct has reached its highest level of type and token frequency and schematic reanalysis, its usage starts to be significantly balanced with creative attempts of re-composition of its internal constituency, with NP increasingly occurring as argument of Pp rather than a first compound member.

Entrenchment and chunking underpin highly automatised patterns of behaviour, such as the way an experienced cook cuts a shallot before making a sauce, or the sequence of steps that an experienced driver gets through when s/he ignites his/her car. Despite the sequence of sensory motor contingencies (cf. Pezzulo 2014: 20) that have been learned whilst performing those actions, people occasionally inhibit the same patterns of behaviour in favour of alternative ways to perform the same task. Individuals are inherently geared towards innovation, which leads to social/linguistic change when a pattern of behaviour is replicated across a community (Croft 2000; Castellano, Fortunato, and Loreto 2007; Traugott and Trousdale 2013). While new patterns of behaviour sometimes affect other members of a social group and thus trigger new social conventions (this in

language change would be a case of reanalysis), on the other hand this study is distinctively focused on the ‘pursuit of innovation’ as such, even when it is restricted to unique instantiations. Entrenchment inhibition thus regard all the cases where individuals ‘divert’ from entrenched behaviour in favour of a less automatised one, which, in turn is designed ‘ad-hoc’ for the here-and-now of their project.

3 The case of NP-Pps

The NP-Pp compounding is often described as a highly productive word formation strategy (Fabb 2001: 68; Plag 2003: 153; Bauer 2006: 490; Bauer et al. 2012: 470). While it is traditionally associated with the passive voice (e.g. Biber et al. 1999: 534, Huddleston and Pullum 2002: 1659, Quirk et al. 1985: 1577), it mostly seems to be characterised by heterogenous argument structure (e.g. Plag 2003). An important aspect of the usage of Np-Pps is that they can be coordinated with adjectives as in *efficient and market-oriented approaches* (cf. Hilpert 2015:117). This is agreed to be due to ambivalent status of participles, which may combine verbal and adjectival features (cf. Huddleston and Pullum 2002: 78).

Bauer et al. (2012: 470), propose a distinction between argumental and non-argumental NP-Pps, the former regarding constructs where the NP can function as the argument of Pp (e.g. in [*drug-related*], *drug* is the prepositional object of *related to drug(s)*), the latter having to do with cases where NP is not a direct argument of Pp (e.g. [*home-cooked*]). They also note that NP-Pp compounds generally seem not to include NP which would collocate as direct objects of the Pp, although some exceptions are acknowledged (Hilpert 2015: 118).

3.1 Are NP-Pps constructions?

What is interesting about the NP-Pp constructs is their relatively recent propagation in American English during the 20th century underpinning both type and token frequency (Ibid.). Figure 1. below illustrates the normalised type frequencies (y axis) of the 20 largest participle families in noun-participle compounding from the COHA (cf. Hilpert 2015: 127):

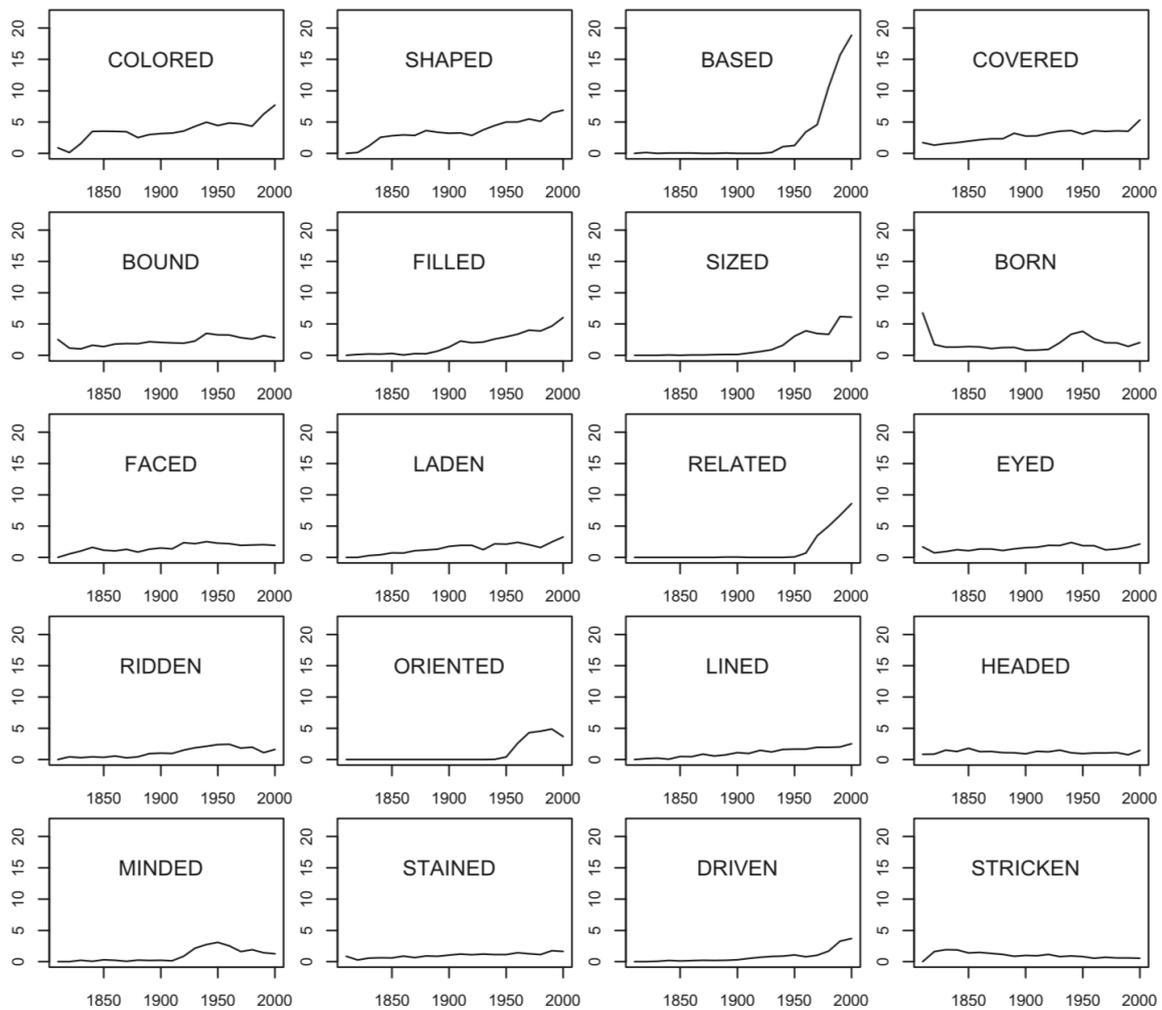


Figure 1.

20 largest participle families in noun-participle type compounding from the COHA (from Hilpert: 2015: 127)

From figure 1 one can notice that certain participles undergo frequency increases starting already in the 19th century (colored, made, shaped, etc.), whilst others show a more recent development, i.e. based, related, sized, oriented. Finally, it is worth acknowledging that only stricken and born undergo decreases (cf. Hilpert 2015: 126). Hilpert (2015) also notes that throughout the 20th century the noun-participle compounding does not include a significant increase of **hapax** legomena (i.e. propagation of the same schema to new components occurring only once) and low-frequency forms and thus can hardly be interpreted as a case of complete constructionalization.

Traugott and Trousdale (2013: 112) address constructionalization based on three well known criteria: schematicity, productivity and compositionality. **Schematicity** underpins ‘routinized, or cognitively entrenched, patterns of experience’ (Kemmer 2003: 78). It involves abstraction across sets of constructions which are closely related to each other in the constructional network (Traugott

and Trousdale 2013: 14). Degrees of schematicity have to do with levels of generality and the extent to which parts of the network are rich in detail (Langacker 2009). In the case of noun-participle compounding, [*drug-related*] would correspond to less schematic node than [NP-*related*], which itself would be less abstract than a higher node [NP-Pp]. **Productivity** impinges on ‘extensibility’ (Barðdal 2008) of a schema to other less schematic constructions and the extent to which this schema is constrained (Boas 2008). Traugott and Trousdale (2013: 209) argue that the pattern [*a lot of X*] constructionalises into a quantifier construction when the X slot accommodates abstract nouns such as *truth* or when the construction refers to plural pronouns (*a lot of sheep* → *they*). After the repeated experience and entrenchment of this and other instances, the schematic representation of the construction is gradually strengthened. Finally, decrease of **compositionality** indicates that the meaning of the whole construction becomes progressively less derivable from the meaning of its parts, e.g. [*believe it or not*] shifting from being an imperative construct (*Believe it or not, as you please, I am decided* - COHA Frou Frou, 1879) to a new intersubjectified parenthetical function (*Then I called back Mrs. Frankenthal and, believe it or not, she was free* - COHA Chairman of the Bored, 1961) (cf. Tantucci 2017a: 113-114). Such reanalysis entails that the imperative mood of the verb *believe* is then less analysable (e.g. it cannot occur in isolation as prototypical imperatives do), together with the meaning of the chunk being comparatively less compositional, no more expressing a transparent command. In the case of NP-Pp compounds, compositionality and analysability underpin the speakers’ possibility to identify the constituents as separate and individual units as they would be when they are not part of the chunk, viz. as elements that can be modified, identified or graded as such: *the street is lined with lovely trees* vs *the street is *luxuriant tree-lined*.

Constructionalization ultimately reflects the emergence of a new node in the constructional network, and can be contrasted with mere constructional change, which denotes “a change affecting one internal dimension of a construction” that “does not involve the creation of a new node” (Traugott and Trousdale 2013: 26).

Hilpert proposes an ‘upward strengthening hypothesis’ to address the distinction between constructs and constructions, namely: “when the experience of a linguistic unit strengthens not only a mental representation of that unit itself, but also a mental representation of a more abstract construction, that process instantiates grammaticalization” (Hilpert 2015: 136). Upward strengthening is thus at play when the mental representation of a construct is not limited to that single unit (e.g. [*Chicago-based*]), but rather ‘climbs up’ to higher node of abstraction (e.g. [NP-*based*] and eventually [NP-Pp]). Hilpert concludes that Noun-participle compounds in the COHA

only remain confined to a limited set of participles and thus fail to ultimately climb up to the more schematic NP-Pp node.

Based on the above, the following analysis has two aims:

1. To assess whether **entrenchment inhibition** significantly acts as a competing force of chunking and constructionalization. This could constitute evidence that chunks can be often reanalysed as compositional combinations, giving rise again to ad-hoc analytic patterns competing with linguistic automatisations.
2. To account for ‘**pre-analysis**’, in the sense of observing whether the internal components of a newly chunked construct used to collocate with one another, viz. whether they had a compositional history prior to the rise of the construct itself (whether the NP *copper* and the Pp *colored* used to collocate syntactically prior to the formation of the NP-Pp *copper-colored*). In sections 5.2 and 5.3 we will propose that formation of NP-Pps without pre-analysis is a powerful diagnostic to assess whether constructionalization is at play. This, in turn, will show that constructionalization may lead to creativity, as people will start to combine the internal constituents of the NP-Pp construction with novel strategies (*shaped in the form of a U - COHA/1946*) that are not found in the first century of the COHA.

4 Constructionalization vs entrenchment inhibition of NP-Pps

This section respectively describes the process of retrieval of our data and provides a distinctive collexeme analysis (DCA) trying to unveil whether entrenchment inhibition significantly intersect with constructional change and possibly constructionalization.

4.1 Data retrieval

To account for the diachronic tension between chunking and entrenchment inhibition of NP-Pps³, we queried the 5 most frequent nominal first-compound-members of each compound with the highest type frequency in the COHA (see fig. 1 from Hilpert (2015) in section 3.1). For instance, in the case of the Pp *based*, we looked for all instances of respectively *Atlanta-based*, *Chicago-based*, *land-based*, *New York-based* and *Washington-based*. We subsequently gathered all the occurrences where the top five most frequent NPs of each compound would collocate within a seven words span

³ This study is focused on NPs (rather than Ns), as the compounds also include phrases that a larger than a single noun, e.g. *A person skiing on a white snow-covered slope will be dark* (COHA/1985).

to the right of their respective Pps, e.g. *based in Chicago, filled with a lot of smoke* and so on (cases where NPs would not be syntactically related to Pps were manually excluded)⁴. This decision aimed at directly tackling both increased schematicity (they are based type-frequencies) but also the actual entrenchment (which relates to tokens, i.a. Bybee 2010) of those top-5 compounds competing with creative attempts of inhibition. At this point, we excluded all specific NP-Pp compounds that did not include a periphrastic alternative expressing a similar in the corpus, e.g. in the COHA there are no periphrastic alternatives to *buisiness-minded* such as **minded towards buisiness*. Similar compounds that were filtered from the analysis were also *base-born, debt-laden, dew-laden, dust-laden, head-lined, stream-lined* and *side-lined*.

Finally, we crucially noticed that a sub-set of Pps does not significantly collocate with their NP as arguments during the 19th century, e.g. [*colored* 7R *cream*]⁵ (7R indicates ‘within a 7 word-span at the right of *colored*’), [*eyed* 7R *goggle*]. This class of NP-Pps did not have a pre-analytical stage before starting to be used as compounds like [*cream-colored*], [*goggle-eyed*]. As it will be discussed in section 5, the presence of this set of Pps is extremely important for assessing whether a process of upward strengthening has been at play. On the other hand, at this stage all Pps without a pre-analytical history were not included in our model. The reason is because this group of NP-Pps have not been through a stage where they were not compounds and thus could not be useful to account for a speaker’s preference for a chunked construct over an alternative compositional expression (e.g. *cream-colored* vs **colored with cream*; *money-eyed* vs **eyed with money*). In table 1 below, the first and second column show all the Pps with a pre-analytical history from our survey, including the 5 most frequent nominal first-compound-members of each compound. In a separate column are also listed all the remaining Pps without pre-analysis, also including their 5 most frequent nominal first-compound-members (fcm).

Pps with pre-analysis	Top 5 nominal fcm	Pps without pre-analysis	Top 5 nominal fcm
<i>based</i>	<i>Atlanta; Chicago; land; New York; Washinton</i>	<i>colored</i>	<i>copper; cream; rose; rust; straw</i>
<i>born</i>	<i>Brooklyn; earth; heaven</i>	<i>eyed</i>	<i>bug; goggle; lynx; money; wall</i>

⁴ This retrieval method is aimed at looking at the combinations of NP-Pp that favoured the most a process of constructionalisation and conventionalisation. A selection of the most frequent NP-Pps among the participle types in figure 1 was methodologically necessary to observe whether entrenchment inhibition is at play as reaction to conventional behaviour. If the survey had counted all the NP-Pps in the COHA, the results would have been less transparent, as they would have included cases that had not yet undergone a clear process of conventionalisation.

⁵ *Cream* appears within 7 word-span to the right of *colored*.

<i>bound</i>	<i>east; leather; south; spell; west</i>	<i>faced</i>	<i>baby; freckle; moon; poker; shame</i>
<i>covered</i>	<i>dust; ivy; moss; snow; vine</i>	<i>headed</i>	<i>gold; level; pig; spear; tow</i>
<i>driven</i>	<i>chauffeur; motor; power; steam; wind</i>	<i>oriented</i>	<i>business; consumer; family; goal; market</i>
<i>filled</i>	<i>gas; smoke; sun; tear; water</i>	<i>ridden</i>	<i>bed; crime; debt; guilt; priest</i>
<i>laden</i>	<i>moisture; snow</i>	<i>shaped</i>	<i>heart; pear; U; V; wedge</i>
<i>lined</i>	<i>fur; head; tree; stream</i>	<i>sized</i>	<i>cap; king; life; man; pint</i>
<i>minded</i>	<i>air; budget; economy; sports</i>		
<i>related</i>	<i>age; church; drug; health; work</i>		
<i>stained</i>	<i>blood; clay; tear; travel; weather</i>		
<i>stricken</i>	<i>grief; horror; panic; poverty; terror</i>		

Table 1.

Most frequent NPs and Pps as NP-Pp compounds both with and without pre-analysis

4.2 Entrenchment inhibition of NP-Pps: A distinctive collexeme account

At this point we annotated all the frequencies of NPs and Pps from the first and the second column in table 1 (section 4.1), both chunked as NP-Pp compounds and syntactically separated. We thus performed a distinctive collexeme analysis (cf. Hilpert, 2006; Gries and Hilpert 2008) measuring the attraction of the lexeme Pp to the NP-Pp construct as opposed to Pp collocating with its NP as a separate argument. Distinctive collexeme analysis (DCA) is often used to compare the distinctive attraction among two (or more) competitive lexemes (or collexemes) with a construction (or collostruct) in different periods of time. In our case, we looked at the same collexeme type Pp (e.g. *based*, *stained*) and we computed the attraction to the NP-Pp collostruct (e.g. [*New York-based*], [*Atlanta-based*]; [*blood-stained*], [*tear-stained*]) while competing with more compositional strategies where all 5 NPs of each Pp would act as a separate argument (e.g. *based in a studio in New York*, *based in Atlanta*; *stained with red blood*, *stained with tears*).

This approach to DCA does not have statistical caveats, as it is based on the same distribution of factors that are included in the contingency table of the Fischer-exact test of collostructional analysis (cf. Hilpert 2006). At the theoretical level, it is yet indeed innovative, as it allows to go beyond the arbitrary selection of the constructions that need to be contrasted with one

another in classic DCA (cf. De Smet 2016 and 2018 on the necessity of a holistic approach to competition in the linguistic system). In fact, this model allows to measure the competition between entrenched vs re-compositional behaviour holistically, without arbitrarily selecting the usage of an individual construction *a* as opposed to another individual construction *b*. The results of our model are based on 8781 annotated occurrences and are given in table 2:

Year	NP-Pp	NP as argument of Pp	Coll strength (logp)
1810	5	7	0
1820	38	51	-0.59
1830	93	134	-2.29
1840	132	258	-11.27
1850	110	205	-8.04
1860	113	189	-5.29
1870	146	239	-7.32
1880	98	225	-12.48
1890	161	233	-4.16
1900	192	221	-2.13
1910	194	226	-1.38
1920	236	241	-1.12
1930	294	218	-0.51
1940	292	180	4.80
1950	309	159	8.87
1960	230	170	1.39
1970	351	151	15.26
1980	492	267	14.23
1990	441	318	3.66
2000	465	331	3.66

Table 2.

Distinctive collexeme analysis of NP-Pps compounding vs NP as argument of Pp

In table 2 above the log-transformed p-values corresponding to each decade measure speakers' preference of Pps occurring in NP-Pp compounds in opposition to more compositional strategies including the same constituents. Positive values indicate a preference for entrenched NP-Pps compounds, while negative ones unveil a distinctive attraction between Pps and NPs acting as separate arguments. The results of this model are plotted in Figure 2:

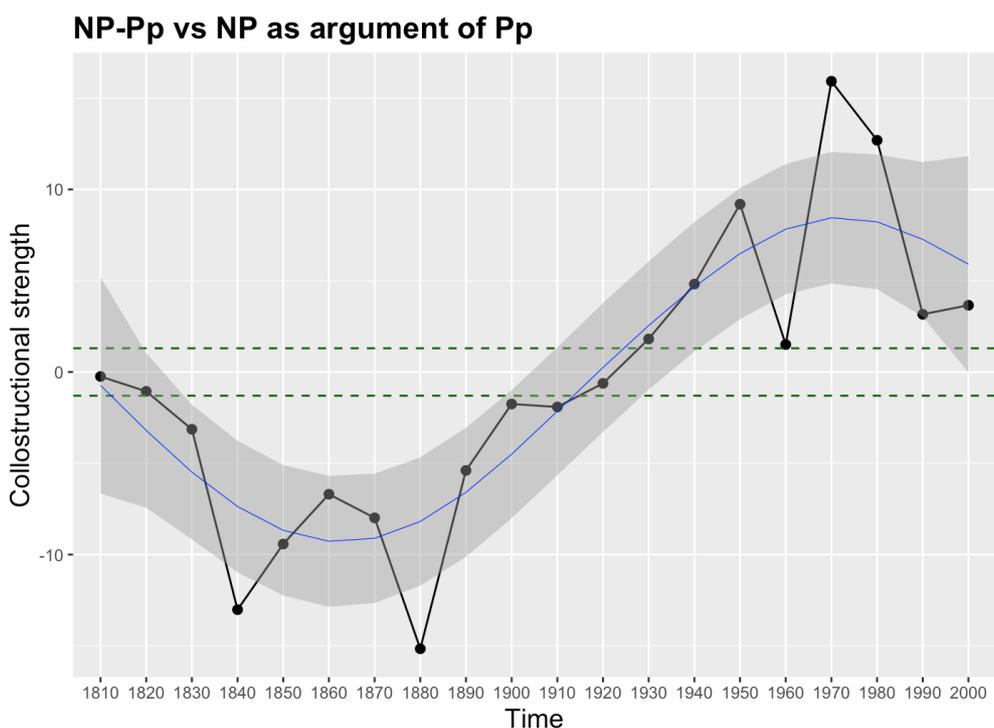


Figure 2.

Distinctive collexeme analysis of NP-Pps competing with NPs as arguments of Pps

Firstly, as expected, the collostructional attraction between Pps and NP-Pp compounds increases over time. In fact, a process of chunking and entrenchment of NP-Pps becomes significant after 1930, as all the remaining observations up to 2000 are above the upper green line intercepting with the y axis at the level of 1.3. The latter corresponds to the minus \log_{10} of 0.05 and can be used as a cut-off point to identify significant collostructional attraction (> 1.3). Conversely, when the values are negative (< -1.3), they indicate a significant preference for the competing option, which in this model underpins NP occurring as a separate argument of Pp. In this second case, significant observations thus fall below the lower green line in figure 2. From this we can note that before the beginning of the 20th century there is a distinctive attraction of Pps to constructs where NPs operate as separate arguments. All in all, the plot confirms what would be predicted in a classic usage-based framework, showing that since 1930 up to the present, a significant process of chunking of NP-Pps is distinctively at play.

That being said, the blue polynomial regression line encompassing the whole period also requires a close inspection. In fact, while all the values after 1930 are undoubtedly above the significance level, it is also important to note a dramatic drop in the speakers' preference for NP-Pps since 1970, with the last two decades barely touching the green line of significance level. This

tendency becomes more interesting after isolating the time-span since one decade before NP-Pp significantly becomes the preferred option:

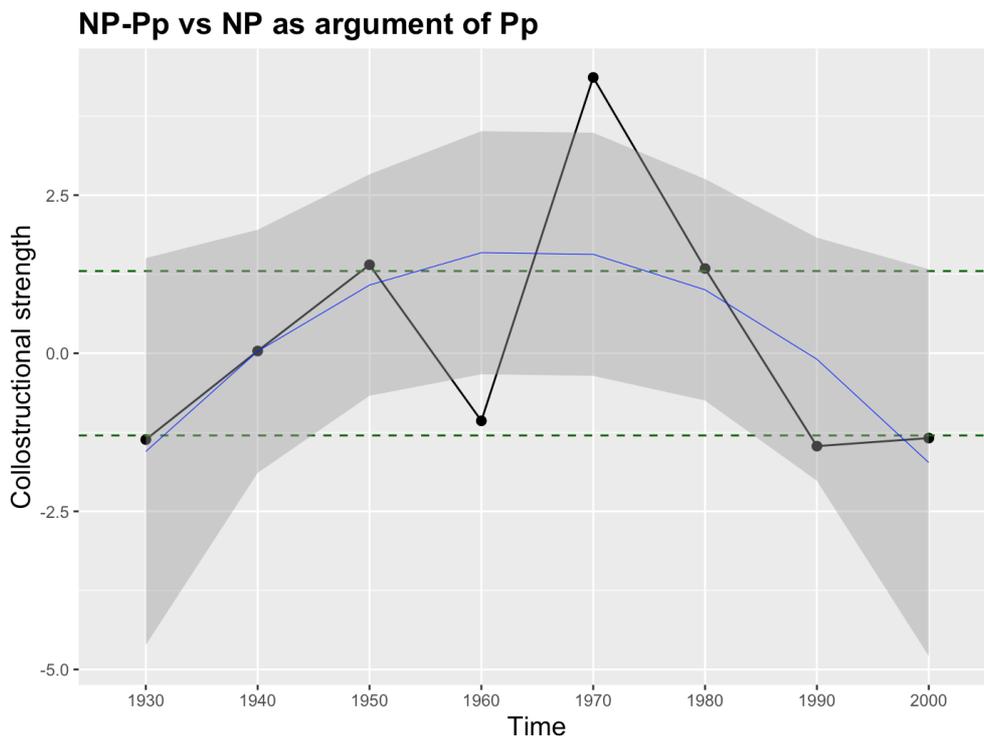


Figure 3.

NP-Pps competing with Entrenchment inhibition from 1930 up to 2000

Figure 3 suggests that after an initial tendency where the use of NP-Pps constructs is increasingly preferred, the collostructional attraction between Pps and their respective entrenched compounds NP-Ps drops from 15.26 in 1970 to 3.66 in 1990 and 2000.

This is somewhat surprising as the increased entrenchment of a relatively new chunked compound is often expected to be exponential and in most cases unidirectional (e.g. Bybee et al. 1994; Hopper 2003; Traugott and Trousdale 2013). As far as we are aware, most usage-based models do not seem to contemplate a competing mechanism that may mitigate on a large-scale an early process of chunking and potential constructionalization (at least not at the operational level of analysis). Yet, the case above clearly indicates that speakers after 1970 eventually refer to the same internal constituents of newly chunked compounds (e.g. [*tree-lined*]) in a more compositional way (*lined on both sides with immense trees*), thus statistically inhibiting the entrenchment of the NP-Pp construct and contributing to ‘drag down’ its collostructional strength from 15.26 to 3.66.

4.3 Entrenchment inhibition as a variable of analysis

The entrenchment inhibition of the NP-Pps with a pre-analytical stage of development can also be captured by plotting their per-million-word (henceforth pmw) frequency together with the pmw frequency of instances where NP collocates as argument of Pp:

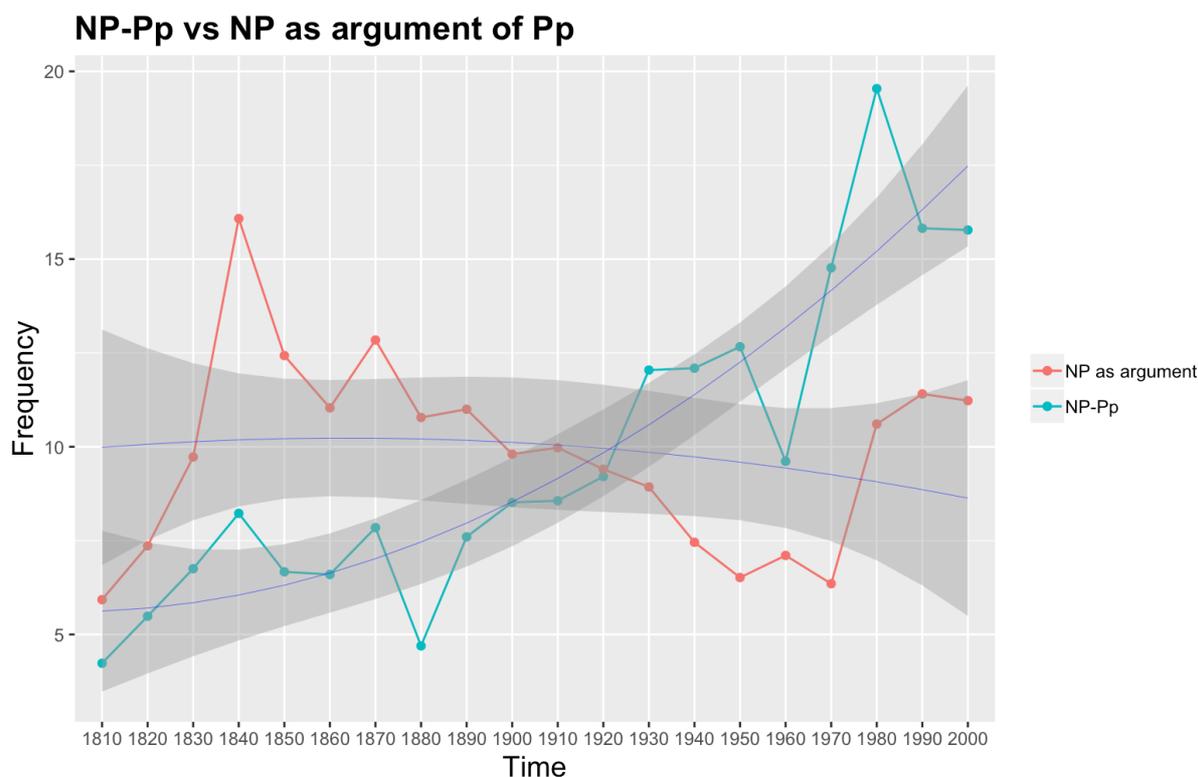


Figure 4.

Normalised frequencies of NP-Pps vs NPs as arguments of Pps

Figure 4 shows that the generalised additive model (GAM) of NP-Pps (light-green line) since 1930 is symmetric to the sharp increase of instances where NP is separated from Pp (red line). Intriguingly, the two lines appear to be almost specular throughout the two centuries of the COHA: the decrease of Np-Pp corresponds to the increase of NP as arguments of Pps and vice versa. Entrenchment inhibition, as an active process of separation of the internal constituents of NP-Pp, can only come into play after 1930, i.e. not before the NP-Pp compounds acquire a significant collostructional attraction to Pp.

The opposite trend of the two polynomial lines is an indicator that entrenchment inhibition may be a decisive variable from a usage-based perspective. To demonstrate this and assess whether the two variables are independent, we used the ‘HoltWinters’ (e.g. Chatfield 1978) function from the R ‘forecast’ package to predict the pmw frequency of NP-Pps compounds during the last 50 years of change (corresponding to the 4th period out of 4 equal time-spans encompassing the

COHA) based on their history from 1810 up to 1950. Notably, we first based our prediction on the development of NP-Pps without accounting for entrenchment inhibition as an additional variable:

Year	Forecast of NP-Pp	Observed NP-Pp	Difference
1960	13.50	9.61	3.89
1970	14.27	14.76	-0.49
1980	15.05	19.54	-4.49
1990	15.82	15.81	-0.01
2000	16.59	15.77	0.82

Table 3.

Forecast of the normalised frequency of NP-Pps from 1950 up to 2000

The mean of the difference between predicted and observed values is -0.04, showing a very accurate forecast of NP-Pp pmw frequency as a single dependent variable of time, with no significant mismatch between predicted and observed frequencies: Kramer 0.092, Chi squared $p > 0.5$ (cf. Cohen 1988 for detailed interpretation of effect sizes):

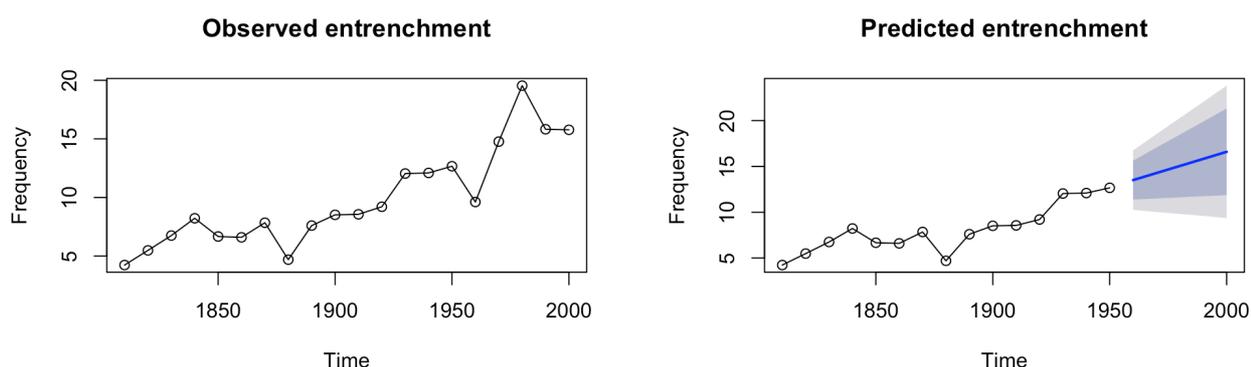


Figure 6.

Observed vs predicted entrenchment of NP-Pps (only based on frequency)⁶

We then similarly plotted a forecast of the last 50 years of development of Pp-Np compounds. However, this time we used the values obtained from the distinctive collexeme analysis in section 4.2, which included entrenchment inhibition as a competing variable:

Year	Forecast of NP-Pp	Observed NP-Pp	Difference
1960	11.13	1.5	9.63
1970	14.13	15.92	-1.79
1980	17.12	12.69	4.43
1990	20.11	3.16	16.19

⁶ The y axis of the two plots has different scaling.

Table 4.

Observed vs predicted entrenchment of NP-Pps (based on collostructional strength)

In this case the mean of the difference between predicted and observed collostructional values is 7.98, which leads to a significant result, Kramer 0.012, Chi-squared $p < 0.0005$:

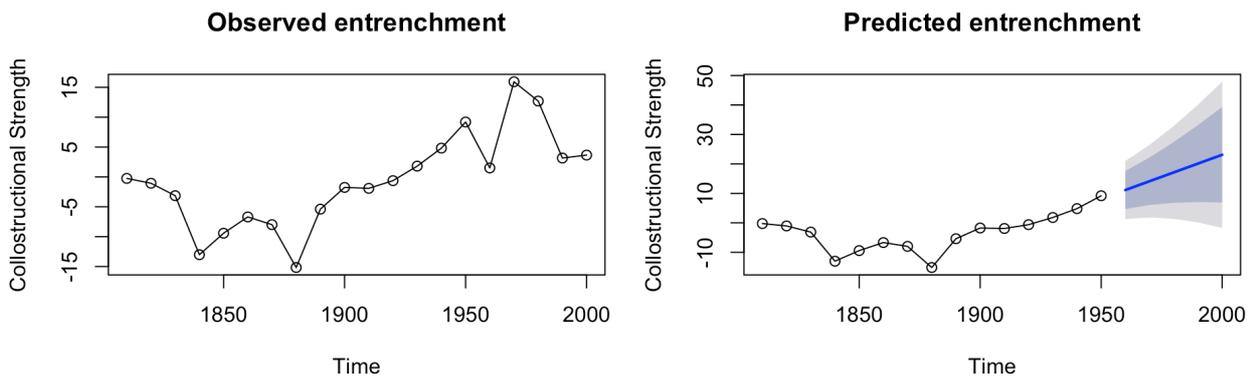


Figure 7.

Observed vs predicted entrenchment of NP-Pps (based on distinctive collexeme values)⁷

The significant mismatch between predicted and observed collostructional strength of NP-Pps crucially suggests that entrenchment inhibition plays a decisive role as a counter-balancing force of chunking and constructional change. In fact, while a trend of constructional change that is only based on frequency can be accurately predicted (see figure 6), things differ when creative strategies involving the same constituents are also taken into account (as in figure 7). In fact, the predicted entrenchment values on the right hand-side plot of figure 7 are significantly higher than what they have been in the last 50 years of the COHA. This significant mismatch sheds light on speakers' inhibition of repetitive usage of those chunked forms in favour of a more compositional surplus of form and meaning. This suggests that entrenchment inhibition is part of the eco-system of constructional change and potentially of online language production. The notion of 'surplus' is a crucial one in research on (im-)politeness (e.g. Kasper 1990; Watts 2003; Tantucci 2018, Tantucci et al. 2018; Tantucci & Wang 2018). In fact, merely formulaic or conventional utterances are sometimes 'enriched' by interlocutors with more than what is expected, thereby implying a greater level of 'ad-hoc' (im-)politeness (hence, the 'surplus' approach). Interestingly, tendencies towards

⁷ The y axis of the two plots has different scaling.

so-called long-term “analyticization” (Haspelmath and Michaelis 2017) are attested of the Romance and Germanic languages, including the Scandinavian (Trudgill 1992: 195–197; Kusters 2003: 231–233, cf. Heine, Kuteva 2005: 32; Gil 2008: 110). This is in line with the idea of paths of degrammaticalization, as they are proposed in Norde (2009). Entrenchment inhibition is somehow connected with the idea that speakers re-discover less reanalysed meanings of a construct(ion). The focus is thus on online choices of momentarily abandoning the usage of automatised form in favour of a more compositional strategy. However, this does not entail that a new long term process of reanalysis has been initiated, simply that an entrenched form has been creatively inhibited.

Importantly, this is not the first case study where a competing mechanism of entrenchment inhibition is being observed diachronically. Tantucci et al. (2018) look at the constructionalization of dialogic pair [A: *good morrow* B | B: (*good*) *morrow* (A)] from the 15th to the 18th century. After reaching the highest degree of entrenchment and automatised form, the dialogic pair is also characterised by a process of re-composition of its internal constituents. In fact, after 1650 the construction shows an increasing tendency to be creatively re-modelled with ad-hoc meanings during online exchanges by means of dynamic resonance (Du Bois, 2014) and non-reciprocal behaviour. This process significantly affected the collostructional attraction of B’s *good morrow* as an entrenched reply to the whole dialogic collostruction [A: *good morrow* B | B: (*good*) *morrow* (A)].

5 Entrenchment inhibition and diachrony

This section aims at providing a specific taxonomy of entrenchment inhibition phenomena. More specifically, we aim to assess whether newly formed chunks without a pre-analytical history (e.g. the NP-Pp [*cream-colored*]) may themselves be subject to entrenchment inhibition. To explain, absence of pre-analysis indicates that the NP (e.g. *cream*) and the Pp (e.g. *colored*) of the compound *cream-colored* did not collocate syntactically before starting to be used as a chunked NP-Pp, as in **colored with cream* or similar. This may entail that a Sp/Wr’s creative attempt of re-composing the internal constituency of a new chunk may not address an immediate node of a constructional network (e.g. [*cream-colored*] or [*X-colored*]), but rather a more schematic one (i.e. [NP-Pp]). This will be discussed as powerful diagnostic of constructionalization.

5.1 Entrenchment inhibition as hyper-, under- and homeo-composition

The analysis in section 4 addresses entrenchment inhibition as a competing variable of constructional change. It is now necessary to assess whether a large-scale process of chunking may in turn affect the perception of the internal constituency of a construct.

To emphasise once more, we argued that entrenchment inhibition occurs ‘on the fly’, as a creative attempt of re-composition of the internal constituency of a conventionalised construction. Crucially, entrenchment inhibition is a large-scale mechanism of creative alteration of repetitive behaviour, and as such does not necessarily lead to innovation or subsequent reanalysis in a linguistic system. Rather, it is inherently part of the eco-system of a conventionalised construction, as it inhibits recurring patterns of verbal behaviour. As a process of re-composition, it may formally differ from pre-analytical strategies involving the same internal constituents. In fact, re-composition may either involve a formal ‘surplus’ or a ‘reduction’ of the elements that used to collocate before a new stage of entrenchment and re-analysis. The former case is what we identify as **hyper-composition**, while the latter can be regarded as **under-composition**. Finally, when the formal structure of re-compositional strategies is formally similar to the pre-analytical behaviour of the same constituents, entrenchment inhibition can be said to underpin **homeo-composition**. In the latter case, the re-discovery of the internal constituency of a construction formally corresponds to the usage of those items before becoming a new chunk (i.e. *Chicago* collocating with *based* in the form of *based in Chicago* despite the new formation of the NP-Pp [*Chicago-based*]). Below we first provide some examples from our dataset including the internal constituents of *bound-south*, respectively being combined in the form of homeo- and under-composition. The NP-Pp compound is given in chevron (< >):

<Side-lined>

homeo-composition: from [*bound to/for the NP*] to [*bound to/for the NP*]

- (1) Our course was south, we knew, for we were **bound to the south pole**.

COHA - Cooper, James Fenimore / *The Monikins* / 1835

- (2) The United States Arsenal in the city, filled with arms and ammunition, was commanded by an officer **bound to the South** [...].

COHA - Robert B. Parker / *A Catskill Eagle* / 1985

under-composition: e.g. from [*bound to/for the NP*] to [*bound NP*]

- (3) The main street was coming to be busy. Along it, together and at intervals, rolled-top buggies, surreys, buckboards, bicycles, hacks and even a work wagon, all **bound south**.

Example (2) above is a case of homeo-composition, as Sp/Wr (Speaker/Writer) inhibits the developing construction [*south-bound*] in favour of the less chunked expression *bound to/for* NP, which itself corresponds to how the NP *south* and the Pp *bound* used to collocate (see (1)), before the arise of the NP-Pp *bound-south*. Conversely, the structure of (3) is quite different. This is a case of under-composition, as this strategy involves a reduced number of items, viz. *bound* NP.

It goes without saying that the meaning that is conveyed through an attempt of re-composition cannot exactly match the one of the NP-Pp construct. This is clearly expected, due to Sp/Wr's marked effort to produce an overt surplus of information that would not be necessary in the case of a conventionalised NP-Pp construct. In examples (4-9) are reported some cases of respectively homeo and hyper composition of the internal constituents of [*blood-stained*] and [*snow-covered*]:

<*Blood-stained*>

homeo-composition: from [*stained with* NP] to [*stained with* NP]

(4) Terrified by phantoms and **stained with blood** shall I not exhibit the tokens of a maniac [...].

COHA - Charles Brockden / The Novels / 1827

(5) The ballot box may be discouragingly slow, but at least it is not **stained with blood**.

COHA - Time Magazine / Two Separatist Strands / 1827

hyper-composition: e.g. from [*stained with* NP] to [*stained AP with* NP POS NP]

(6) I still had the dollars I'd come with, **stained brown with my blood** but no less negotiable.

COHA - E. L. Doctorow / Look Lake / 1980

<*Snow-covered*>

homeo-composition: [from *covered with* NP] to [*covered with* NP]

(7) [...] and we had a distant view of part of the Andes, which appeared **covered with snow**.

COHA - Journal of a Cruise to the Pacifick Ocean/ North American Review/ 1815

(8) The river looked dark and clean, its frozen banks were **covered with snow**.

COHA - Friedel Ungeheuer / Return to Frankfurt / 1970

hyper-composition: e.g. from [*covered with* NP] to [*covered in* NP *of* NP]

- (9) As our wagon moved slowly past fields **covered in** deep drifts **of** encrusted **snow**, I looked expectantly about for farm houses [...].

COHA - Kent, Kathleen / The heretic's daughter / 2000

5.2 The re-composition of NP-Pps without pre-analysis

This section aims at enquiring the re-compositional strategies of NP-Pps without a pre-analytical stage of usage, listed again in table 6:

Pps without pre-analysis	Top 5 nominal fcm
<i>colored</i>	<i>copper; cream; rose; rust; straw</i>
<i>eyed</i>	<i>bug; google; lynx; money; wall</i>
<i>faced</i>	<i>baby; freckle; moon; poker; shame</i>
<i>headed</i>	<i>gold; level; pig; spear; tow</i>
<i>oriented</i>	<i>buisness; consumer; family; goal; market</i>
<i>ridden</i>	<i>bed; crime; debt; guilt; priest</i>
<i>shaped</i>	<i>heart; pear; U; V; wedge</i>
<i>sized</i>	<i>cap; king; life; man; pint</i>

Table 6.

Pps and their respective NPs without pre-analysis

Entrenchment inhibition as such may only occur after a process of chunking of NP-Pp started to occur. This entails that a formal process of re-composition can only be at stake starting from the beginning of 1900. Predictably, frequencies of cases where speakers re-compose constructs that were almost not existent before being used as NP-Pps, are relatively low. However, such constructs without a pre-analytical stage crucially show a significant increase of specifically hyper-compositional strategies during the 20th century. The left-hand plot in figure 9 below reports the log-transformed frequencies of re-composition phenomena in hundred-million words (phmw) from the first to the second half of the 20th century:

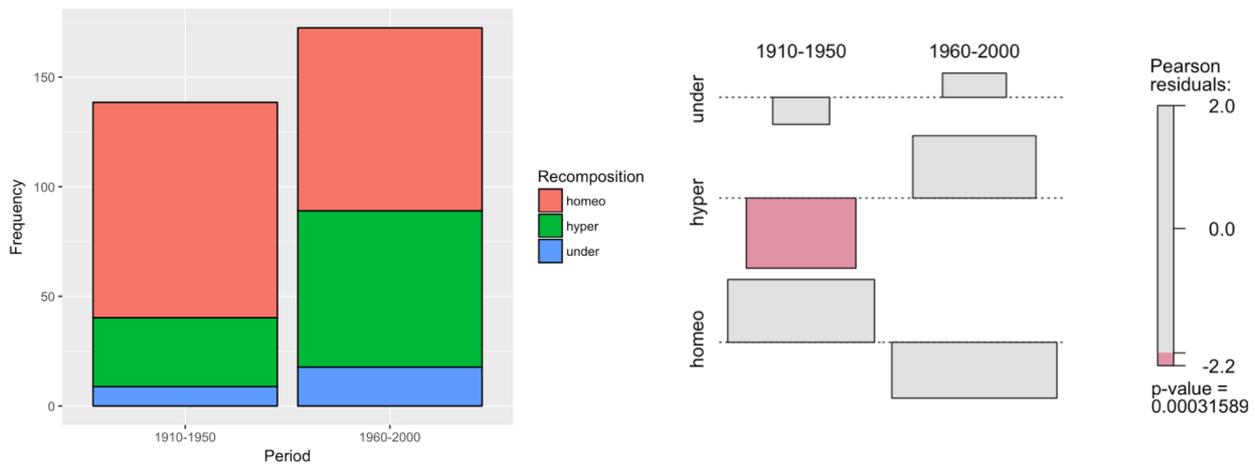


Figure 8.
Re-composition NP-Pps without pre-analysis

The right-hand side of figure 8 illustrates the Pearson residuals of predicted and observed frequencies (negative residuals appear in red). From the plot, it clearly emerges a significant increase of hyper-compositional strategies during the 20th century. This tendency may indeed be connected to the fact that all combinations from table 6 did not include a pre-analytical stage of usage before the formation of NP-Pp compounds (i.e. before the 20th century they did not collocate used as separate items). Speakers may thus creatively re-compose the internal constituents of the construct without having in mind pre-existing idiomatic combinations of Pps and their respective NPs such as **colored with copper* or **shaped with/of pear(s)*. Simply put, the significant increase of hyper-composition during the second half of the 20th century seems to directly correlate with the progressive formation of NP-Pp as a new schematic construction. It is specifically after the formation of the NP-Pp node, that speakers tend to distinctively hyper-compose meanings with novel strategies and creative forms that were not in use during the 19th century. Some cases of hyper-composition of Np-Pps without pre-analysis are given below:

<U-shaped>

- (13) The gaseous diffusion plant for separating U.235 is **shaped** in the form of a **U** and covers an area of several million square feet.

COHA - William L. Lawrence / Dawn Over Zero / 1946

<Rose-colored>

- (14) It was a rock-crystal, **colored** faintly deep within with amethyst and **rose**, but clear as water.

COHA - Victoria Holt / On The Night of The Seventh Moon / 1972

<Shame-faced>

- (15) Each glanced pleasantly at the other's medal. They **faced** each other without **shame**. Neither had the slightest sense of hypocrisy either in himself [...].

COHA – Variou Tantuccis / Short Stories of Various Types / 1920

<Consumer-oriented>

- (16) Album, however, seems **oriented** not so much to casual **consumers** as to users with a consuming passion for digital [...].

COHA – Henry Norr / Programs help you get the picture / 2003

The occurrence of new periphrastic usages with the internal constituents of NP-Pps without pre-analysis entails that constructional change as such may, in turn, lead to creative behaviour. In fact, new periphrastic forms as in (13-16) are the result of the propagation of the NP-Pp schema to items that previously did not collocate with one another.

One last remark needs to be made about genre, which tends to be constant throughout the two centuries encompassing the COHA, including respectively “Fiction”, “Popular Magazines”, “Newspapers” and “Non-fiction”. The only exception is represented by the “Newspapers” section, which is absent from 1810 until 1850. This is however a genre where creative variations of idiomatic constructs are intuitively going to be less frequent in comparison with the “Fiction” section. In this sense, we must remark that the first quarter of the COHA includes comparatively more fictional data. This may be a fair reason to expect the first quarter of the corpus to include a higher proportion of instances of re-composition. However, both surveys from this study have shown that re-composition grows during the 20th century either in the form of under- or homeo-composition regardless of the genre where they appear.

5.3 Absence of pre-analysis as a diagnostic of schematicity and upward strengthening

The significant development of hyper-compositional strategies of items without a pre-analytical history is a useful diagnostic for assessing whether NP-Pp has been through a process of constructionalization, thus whether a new node has emerged in the constructional network. Intuitively, a process of upward strengthening reaching a schematic NP-Pp node is a reasonable explanation for the propagation of new compounds such as *heart-shaped* or *pint-sized*. To explain, while some of the Pps from table 6 may collocate in the COHA in compounds with adjectives in the

place of the NP slot (e.g. *bright-colored*, *sweet-faced* or *clear-eyed*), other participles from the same dataset such as *sized*, *shaped*, *oriented* and others can only collocate with a NP as their first compound member. This entails that their propagation as compound members can only originate from a [NP-Pp] node, rather than being mapped from less schematic nodes entrenched ‘locally’ in the form of [X-colored], [X-faced] or [X-eyed].

In Hilpert (2015) it is noted that **hapax** legomena do not arise significantly after both the type and the token increase of the NP-Pp construct. This is posited as one of the reasons to reject an upward strengthening process being fully at play throughout the development of the compound. Yet, he also crucially suggests that upward strengthening is to be intended as a gradual phenomenon which implicitly validates the idea that the missed increase of hapax instances may not be the sole diagnostic to validate the hypothesis applied to the NP-Pp schema. Notably, the consistent increase of Np-Pps that lack a more compositional counterpart before the end of the 19th century cannot be overlooked. This is an indicator of productivity (cf. Traugott and Trousdale 2013) occurring directly from a higher NP-Pp node rather than mere entrenchment of two words occurring together.

Even if this compounding strategy is limited to a (yet still relatively large) set of items, the extension of the Np-Pp configuration to constructs without a compositional alternative can help to assess whether upward strengthening has been at play. In this sense, the presence of pre-analysis can be adopted as a corpus-driven diagnostic to cast light on the degree of schematicity of the node allowing for the context-expansion (Himmelman 2004) and analogization (Traugott and Trousdale 2013) of a construction.

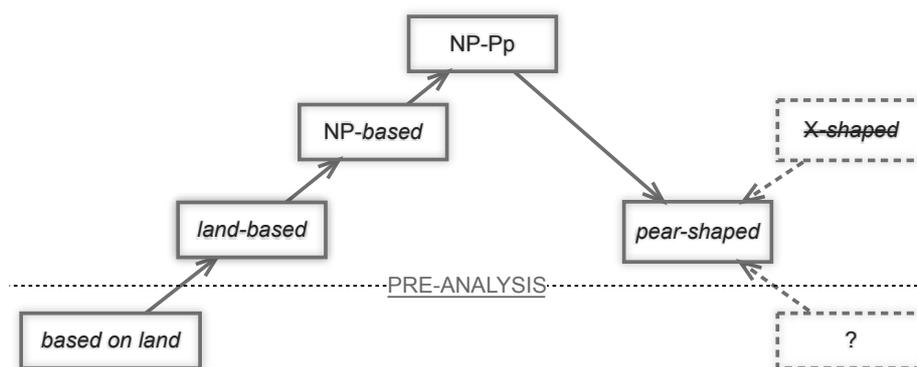


Figure 9
Schematicity, pre-analysis and NP-Pps

The left-hand side of figure 9 illustrates the upward-strengthening process of NP-Pps with a pre-analytical history (e.g. *based on land*) until contributing to reaching the highest node of

schematicity. On the right-hand side of the diagram it is shown how the extension to other participles such as in *pear-shaped*, must indeed originate directly from the NP-Pp node (e.g. not from alternative [ADJ-Pp], [ADV-Pp] ones). This is crucially due to the fact that NP-Pps of the kind of *pear-shaped* do not include pre-analytical strategies, nor their Pp occurs in compounds with first compound members that are not NPs⁸.

The issue as to whether NP-Pps are to be considered as fully schematic constructions is still open to further analysis, as pragmatic constraints may also be at play in the formation of specific compounds, e.g. the Pp *shaped* requiring a NP profiling a simple and schematic image, such as *pear*, *U* or *V* (cf. Goldberg and Ackerman 2001: 811 on this specific issue). The main argument of this section is that pre-analysis can be used a powerful criterion for assessing whether an on-going process of up-ward strengthening, increasing schematicity and constructionalization is at play.

6 Conclusions

In this study we proposed to account for constructional change as a mechanism that is not independent from large-scale competitive attempts of entrenchment inhibition. We have argued that the constructional change of NP-Pps is significantly affected by speakers' creative attempts to re-compose the internal constituency of the compounds. This phenomenon is confirmed by the fact that a subset of NP-Pps without a compositional history prior to the formation of the compound, starts to be re-combined with completely novel strategies during the second half of the 20th century. What is crucial of the present account, is that strategies of entrenchment inhibition appear to be part of the 'ecosystem' of an established construction, viz. it is often the case that speakers creatively re-compose the internal meaning of a chunk without necessarily triggering further stages of innovation and reanalysis in the linguistic system. Entrenchment inhibition can thus be identified as a creative large-scale mechanism that inherently **counter-balances** the conventionalisation and the repetitive behaviour of entrenched linguistic forms. We provide a taxonomy of re-compositional phenomena and formally introduce the notion of pre-analysis as a complementary corpus-driven diagnostic to

⁸ While the above is an ideal re-construction of the the relationship between absence of pre-analysis and productivity, a variety of possible source constructions might be clearly be involved in the same process. As suggested by one reviewer, Old English had noun incorporation and some NP-Pp compounds may have emerged as the past participles of compound verbs (e.g. *goldwrecan* 'inlay with gold'). In addition, Noun-verb conversion was already productive in early Modern English (e.g. the Oxford English Dictionary s.v. *nosed* gives *She ys myche lyke nosid vnto the quyn hir moder*), implying that compounds like *hawknosed* may have been built on once existent verbs. While this would be worth closer investigation, it is beyond the scope of the current paper.

assess whether a construct is undergoing a process of upward strengthening, increased schematicity and thus constructionalization.

All in all, we propose that creativity should not only be addressed as an exceptional mechanism of extravagant innovation and subsequent reanalysis. Crucially, creativity is also a very common process of mitigation of conventionalised and repetitive behaviour. In this sense, this study additionally aims to address research in cognitive architectures and artificial intelligence by problematising creativity as a recurrent mechanism ‘inhibiting’ conventional chunks of behaviour and not simply a problem-solving one.

Bibliography

- Alloppenna, Paul D, James S Magnuson, and Michael K Tanenhaus. 1998. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language* 38 (4): 419-439.
- Ambridge, Ben, Julian M Pine, Caroline F Rowland, and Chris R Young. 2008. The effect of verb semantic class and verb frequency (entrenchment) on children's and adults' graded judgements of argument-structure overgeneralization errors. *Cognition* 106 (1): 87-129.
- Andersen, H. 2001. Actualization and the (uni)directionality. In H. Andersen (Ed.), *Actualization: Linguistic change in progress* (Vol. 219, pp. 225–248). Amsterdam: John Benjamins.
- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., and Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111(4), 1036.
- Anderson, M. C., Bjork, R. A., and Bjork, E. L. 1994. Remembering can cause forgetting: retrieval dynamics in long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20(5), 1063.
- Anderson, M. C., and Neely, J. H. 1996. Interference and inhibition in memory retrieval. *Memory* (pp. 237-313): Elsevier.
- Arnold J. E., T. Wasow, A. Losongco and R. Ginstrom. 2000. Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language* 76. 28–55.
- Baker, C. L. 1979. Syntactic theory and the projection problem. *Linguistic Inquiry* 10(4). 533–581.
- Barðdal, Jóhanna. 2008. *Productivity. Evidence from case and argument structure in Icelandic*. Amsterdam: John Benjamin Publishing Company.
- Bauer, Laurie. 2006. Compounds and minor word-formation types. In Aarts, Bas, and April M. S. McMahon, eds. 2006. *The handbook of english linguistics*. Blackwell Handbooks in Linguistics. Malden, MA ; Oxford: Blackwell Pub, 483-506.
- Bauer, Laurie, Rochelle Lieber and Ingo Plag. 2012. *The Oxford reference guide to English morphology*. Oxford: Oxford University Press.

- Beckner, Clay, Richard Blythe, Joan Bybee, Morten H Christiansen, William Croft, Nick C Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman, and Tom Schoenemann. 2009. Language is a complex adaptive system: Position paper. *Language Learning* 59 (s1): 1-26.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman grammar of spoken and written english*. London: Longman.
- Boas, Hans C. 2008. Resolving form-meaning discrepancies. In Jaako Leino, ed., *Constructional Reorganization*, 11–36. Amsterdam: Benjamins.
- Boyd, Jeremy K, and Adele E Goldberg. 2011. Learning what not to say: The role of statistical preemption and categorization in a-adjective production. *Language* 87 (1): 55-83.
- Bowerman, M. 1988. The ‘no negative evidence’ problem: How do children avoid constructing an over general grammar? In J. A. Hawkins (ed.), *Explaining language universals*, 73–101. Oxford: Basil Blackwell.
- Braine, M. D. S. and P. J. Brooks. 1995. Verb argument structure and the problem of avoiding an overgeneral grammar. In M. Tomasello and W. E. Merriman (eds.), *Beyond names for things: Young children’s acquisition of verbs*, 353–376. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Brown, R. and C. Hanlon. 1970. Derivational complexity and order of acquisition in child speech. In J. R. Hayes (ed.), *Cognition and the development of language*. New York, NY: Wiley.
- Bybee, J. 1998. The emergent lexicon. In M. C. Gruber, D. Higgins, K. S. Olson, and T. Wysocki (Eds.), *Papers from the thirty-fourth regional meeting of the Chicago linguistic society*, 421–435. Chicago: Chicago Linguistics Society.
- Bybee, J. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change* 14 (3): 261-290.
- Bybee, J. 2003. *Phonology and language use*. Vol. 94. Cambridge: Cambridge University Press.
- Bybee, J. 2010. *Language, usage and cognition*. Cambridge: Cambridge University Press.
- Bybee, Joan L, Revere Dale Perkins, and William Pagliuca. 1994. *The evolution of grammar: tense, aspect, and modality in the languages of the world*. Vol. 196. Chicago: University of Chicago Press.
- Chatfield, Chris. 1978. 'The holt-winters forecasting procedure'. *Applied Statistics*, 264-279.
- Clark, E. 1987. The principle of contrast: A constraint on language acquisition. In B. MacWhinney (ed.), *Mechanisms of language acquisition*, 1–33. Hillsdale, NJ: Erlbaum.
- Croft, W. 2000. *Explaining language change: An evolutionary approach*. London: Longman.
- Croft, W. 2010. The origins of grammaticalization in the verbalization of experience. *Linguistics* 48 (1): 1-48.
- Croft, William, and D Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.
- Davis, Mark. 2012. Expanding horizons in historical linguistics with the 400 million word corpus of historical American English”. *Corpora* 7: 121-57.
- De Smet, H., D’hoedt, F., Fonteyn, L., and Van Goethem, K. (2018). The changing functions of competing forms: Attraction and differentiation. *Cognitive Linguistics*, 29(2), 197-234.
- Diessel, Holger. 2011. "Grammaticalization and language acquisition." In *The Oxford handbook of grammaticalization*, edited by Heiko Narrog and Bernd Heine, 130-141. Oxford: Oxford University Press.
- Di Sciullo, Anna-Maria, and Edwin Williams. 1987. *On the definition of word*. Vol. 14. Cambridge, Mass.: MIT Press.
- Du Bois, John W. 2014. Towards a dialogic syntax. *Cognitive Linguistics* 25 (3): 359-410. <https://doi.org/10.1515/cog-2014-0024>.

- Ellis, Nick C, and Diane Larsen-Freeman. 2006. "Language emergence: Implications for applied linguistics—Introduction to the special issue." *Applied linguistics* 27 (4):558-589.
- Ellis, Nick C. 2017. Chunking in language usage, learning, and change: I don't know. In: Hundt, Marianne, Mollin, Sandra, Pfenninger, Simone (Eds.), *The Changing English Language. Psycholinguistic Perspectives*. Cambridge University Press, Cambridge.
- Fabb, Nigel. 2001. Weak monosyllables in iambic verse and the communication of metrical form. *Lingua* 111 (11): 771-790.
- Foraker, S., T. Regier, N. Khetarpal, A. Perfors, and J. B. Tenenbaum. 2007. Indirect evidence and the poverty of the stimulus: The case of anaphoric one. In D. S. McNamara and J. G. Trafton (eds.), *Proceedings of the Twenty-Ninth Annual Cognitive Science Society*, 275–80. New York, NY: Lawrence Erlbaum Associates.
- Gennari, Silvia P, and Maryellen C MacDonald. 2008. Semantic indeterminacy in object relative clauses. *Journal of Memory and Language* 58 (2): 161-187.
- Gibson, E., L. Bergen and S. T. Piantadosi. 2013. Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. *Proceedings of the National Academy of Sciences* 110(20). 8051–8056.
- Goldberg, Adele E. 1993. Another look at some learnability paradoxes. In *The proceedings of the twenty-fifth annual child language research forum*, EV Clark (ed.), 60-75.
- Goldberg, Adele E. 1995. *Constructions. A construction grammar approach to argument structure*. Chicago: University of Chicago Press.
- Goldberg, Adele E. 2006. *Constructions at work*. Oxford: Oxford University Press.
- Goldberg, Adele E. 2011. Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics* 22 (1): 131-153.
- Goldberg, A. E., and Ackerman, F. (2001). The pragmatics of obligatory adjuncts. *Language*, 798-814.
- Gregory, M., Raymond, W., Bell, A., Fosler-Lussier, E., and Jurafsky, D. (1999). The effects of collocational strength and contextual predictability in lexical production. Paper presented at the Chicago Linguistic Society, Chicago.
- Gries, Stefan Th., and Martin Hilpert. 2008. 'The identification of stages in diachronic data: Variability-based neighbour clustering'. *Corpora* 3 (1): 59-81. <https://doi.org/10.3366/E1749503208000075>.
- Gries, S. T., Hampe, B., and Schönefeld, D. 2010. Converging evidence: more on the association of verbs and constructions. In S. Rice and J. Newman (Eds.), *Empirical and experimental methods in cognitive/functional research* (Vol. II, pp. 59–72). Stanford, CA: CSLI.
- Haiman, J. 1994. Ritualization and the development of language. In W. Pagliuca (Ed.), *Perspectives on grammaticalization* (pp. 3–28). Amsterdam/Philadelphia: John Benjamins.
- Haspelmath, M. 1999. Why is grammaticalization irreversible? *Linguistics*, 37(6), 1043–1068.
- Haspelmath, Martin and Susanne Maria Michaelis. (2017). Analytic and synthetic: Typological change in varieties of European languages in Language Variation. In Isabelle Buchstaller and Beat Siebenhaar (eds.), *European Perspectives VI: Selected Papers from the 8th International Conference on Language Variation in Europe (ICLaVE 8), Leipzig 2015*, 3–22. Amsterdam: Benjamins.
- Heine, B., and T. Kuteva. 2007. *The Genesis of grammar: A reconstruction*. Vol. 9. Oxford: Oxford University Press.
- Hilpert, M. 2006. Distinctive collexeme analysis and diachrony. *Corpus linguistics and linguistic theory* 2 (2). <https://doi.org/10.1515/CLLT.2006.012>.

- Hilpert, M. 2015. From hand-carved to computer-based: noun-participle compounding and the upward strengthening hypothesis. *Cognitive Linguistics* 26 (1). <https://doi.org/10.1515/cog-2014-0001>.
- Himmelman, N. P. 2004. Lexicalization and grammaticization: Opposite or orthogonal. In W. Bisang, N. P. Himmelman, and B. Wiemer (eds). *What makes grammaticalization?: A look from its fringes and its components*, 21-41. Vol. 158. Berlin: Walter de Gruyter.
- Hopper, P. J., and Traugott, E. C. (2003). *Grammaticalization* (2nd ed.). Cambridge: Cambridge University Press.
- Huddleston, R. and G. K. Pullum. 2002. *The Cambridge grammar of English*. Cambridge: Cambridge University Press.
- Kamide, Y., G. Altmann and S. L. Haywood. 2003. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language* 49(1). 133–156.
- Kemmer, S. 2003. Schemas and lexical blends. In H. Cuyckens, T. Berg, R. Dirven, and K. Panther, (eds). *Motivation in language*, 69-97. Amsterdam: John Benjamin Publishing Company.
- Lakoff, George. 1970. *Irregularity in syntax*. New York, NY: Holt, Rinehart and Winston.
- Langacker, R. W. 2009. *Investigations in cognitive grammar*. Vol. 42. Berlin: Walter de Gruyter.
- Levshina, N. 2015. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/z.195>.
- Jaeger, T. F. and N. E. Snider. 2013. Alignment as a consequence of expectation adaptation: syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition* 127 (1): 57-83.
- Jurafsky, D., Bell, A., Gregory, M., and Raymond, W. 2001. Probabilistic relations between words: evidence from reduction in lexical production. In J. Bybee and P. Hopper (Eds.), *Frequency and the emergence of linguistic structure* (pp. 229–254). Amsterdam/Philadelphia: Benjamins.
- Jurafsky, D., Bell, A., and Girand, C. 2002. *The role of the lemma in form variation*. Paper presented at the Papers in laboratory phonology, Berlin/New York.
- Kasper, G. 1990. Linguistic politeness:: Current research issues. *Journal of Pragmatics*, 14(2), 193-218.
- Langacker, R. W. 1987. *Foundations of cognitive grammar: Theoretical prerequisites* (Vol. I). Stanford, CA: Stanford University Press.
- McRae, K., de Sa, V. R., and Seidenberg, M. S. 1997. On the nature and scope of featural representations of word meaning. *Journal of Experimental Psychology: General*, 126(2), 99.
- Newell, A. 1990. *Unified theories of cognition*. Cambridge, MA: MIT Press.
- Norde, M. 2009. *Degrammaticalization*. Oxford: Oxford University Press.
- Norman, K. A., Newman, E. L., and Detre, G. 2007. A neural network model of retrieval-induced forgetting. *Psychological Review*, 114(4), 887.
- Oppenheim, G. M., Dell, G. S., and Schwartz, M. F. (2010). The dark side of incremental learning: A model of cumulative semantic interference during lexical access in speech production. *Cognition*, 114(2), 227-252.
- Pinker, S. 1981. On the acquisition of grammatical morphemes. *Journal of Child Language* 8 (2): 477-484.
- Plag, I. 2003. *Word-formation in English*. Cambridge: Cambridge University Press.
- Quirk, R., S. Greenbaum, G. Leech, and J. Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Roland, D., Yun, H., Koenig, J.-P., and Mauener, G. 2012. Semantic similarity, predictability, and models of sentence processing. *Cognition*, 122(3), 267-279.

- Robenalt, C. and A. E. Goldberg. 2015. Judgment evidence for statistical preemption: It is relatively better to vanish than to disappear a rabbit, but a lifeguard can equally well backstroke or swim children to shore. *Cognitive Linguistics* 26 (3): 467-503.
- Searle, J. R. 1969. *Speech acts: an essay in the philosophy of language*. Vol. 626. Cambridge: Cambridge University Press.
- Schmid, H.-J. 2007. Entrenchment, salience, and basic levels. *The Oxford handbook of cognitive linguistics*, 117138.
- Schmid, H.-J. 2017. A framework for understanding linguistic entrenchment and its psychological foundations. In *Entrenchment and the psychology of language learning: how we reorganize and adapt linguistic knowledge*, (ed.) Hans-Jörg Schmid, 9–38. Washington, D.C.: American Psychological Association; Berlin: Walter de Gruyter GmbH.
- Sinclair, J. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.
- Tagliamonte, S. A. and R. H. Baayen. 2012. Models, forests, and trees of York English: was/were variation as a case study for statistical practice. *Language Variation and Change* 24 (2): 135-178.
- Tantucci, V. 2013. Interpersonal Evidentiality: The Mandarin V-过 guo construction and other evidential systems beyond the ‘source of information’. *Journal of Pragmatics*, 57, 210–230.
- Tantucci, V. 2015. Epistemic inclination and factualization: a synchronic and diachronic study on the semantic gradience of factuality. *Language and Cognition*, 7(3), 371-414.
- Tantucci, V. 2017a. From immediate to extended intersubjectification: A gradient approach to intersubjective awareness and semasiological change. *Language and Cognition*, 9(1), 88–120.
- Tantucci, V. 2017b. An evolutionary approach to semasiological change: Overt influence attempts through the development of the Mandarin 吧-ba particle. *Journal of Pragmatics*, 120, 35–53.
- Tantucci, V. 2018. From co-actionality to extended intersubjectivity: Drawing on language change and ontogenetic development. *Applied linguistics*. doi:<https://doi.org/10.1093/applin/amy050>.
- Tantucci, V., Culpeper, J., & Di Cristofaro, M. (2018). Dynamic resonance and social reciprocity in language change: the case of Good morrow. *Language Sciences*, 68, 6-21. doi:<https://doi.org/10.1016/j.langsci.2017.09.004>
- Tantucci, V., and Wang, A. Q. 2018. Illocutional concurrences: The case of evaluative speech acts and face-work in spoken Mandarin and American English. *Journal of Pragmatics*(138), 60–76.
- Terkourafi, M. 2015. Conventionalization: A new agenda for im/politeness research. *Journal of Pragmatics* 86: 11-18.
- Traugott, E. C., and Dasher, R. B. 2002. *Regularity in semantic change*. Cambridge: Cambridge University Press.
- Traugott, E. C., and Trousdale, G. 2013. *Constructionalization and constructional changes*. Oxford: Oxford University Press.
- Trude, A., and Nozari, N. 2017. *Inhibitory Control Supports Referential Context Use in Language Production and Comprehension*. Paper presented at the CogSci.
- Watts, R. J. (2003). *Politeness*: Cambridge University Press.
- Wood, W., and Neal, D. T. 2007. A new look at habits and the habit-goal interface. *Psychological Review*, 114(4), 843.

