Gilquin, G. 2022. 'Constructing learner speech: On the use of spoken data in Applied Construction Grammar'. In H. Boas (ed.) *Directions for Pedagogical Construction Grammar: Learning and Teaching (with) Constructions* (pp. 73-96). Berlin: de Gruyter.

# Constructing learner speech: On the use of spoken data in Applied Construction Grammar

Gaëtanelle Gilquin
Université catholique de Louvain

**Abstract**
Construction Grammar, like many other theoretical frameworks in linguistics, tends to be characterized by a bias toward the study of written, rather than spoken, language. This bias is also visible in the young field of Applied Construction Grammar, which is the focus of this paper. Adopting a three-fold perspective, the paper pleads for more research on speech in Applied Construction Grammar. First, it argues that, from a descriptive point of view, we cannot dispense with the investigation of constructions (and constructicons) in speech and in learner speech in particular. Second, turning to applications among foreign language learners, it highlights the potential benefits of the study of spoken language for the learning and teaching of constructions. Finally, it considers more theoretical issues, showing how the recognition of speech can help refine the (Applied) Construction Grammar model.

## 1 Introduction

The Latin proverb according to which spoken words fly away and written words remain (*verba volant, scripta manent*) emphasizes the ephemeral nature of speech. Thanks to technological advances, it has become possible to capture speech for linguistic purposes, most notably in the form of spoken corpora, and to make it permanent, just like writing. Once transcribed, spoken language can be analysed in a systematic way, using the same tools as those developed for the linguistic analysis of writing. As for the acoustic signal, it can be studied by means of special linguistic programs, which have been specifically designed to deal with audio files. Yet, for reasons that may range from the very practical (difficulty of collecting and analysing spoken data) to the very ideological (lack of consideration for spoken forms of communication), speech has, on the whole, received less scholarly attention than writing. While this is true of many linguistic frameworks, we will focus here on Construction Grammar, and more particularly Applied Construction Grammar, which is concerned with "the acquisitional and pedagogical implications of Construction Grammar in second/foreign language teaching and learning" (Gilquin and De Knop 2016: 8; see also the contributions in De Knop and Gilquin 2016).

Starting from the observation that the study of speech has been neglected in (Applied) Construction Grammar (Section 2), this paper pleads for more research on spoken language data in this framework, arguing that it can be beneficial from a descriptive, applied and theoretical perspective. From a descriptive point of view, we will consider the comparison of spoken and written constructions (Section 3.1), as well as the study of the spoken learner constructicon (Section 3.2). From a more directly applied point of view, we will see how spoken data can be used to better understand foreign language learning (Section 4.1) and to improve foreign language teaching (Section 4.2). From a theoretical point of view, it will be shown that the study of speech can help refine the Construction Grammar model (Section 5).

This overview of the benefits that Applied Construction Grammar can derive from spoken language research relies on the discussion of relevant works taken from the literature. Not all of these works are embedded within the theories of Construction Grammar, but they are all deemed compatible with (some aspects of) these theories, and in particular with the tenets of Goldbergian Construction Grammar. Similarly, the studies that are cited are not

necessarily related to learner language, but they are applicable to it. Most of them have to do with the English language, but the general principles emerging from these studies are believed to be potentially valid for other languages as well. It should also be pointed out that the perspective that is adopted in this paper is that of a foreign (not second) language learning context. As briefly shown in Section 4.1, the situation for second language learners may differ in certain respects.

## 2 A written bias in (Applied) Construction Grammar

Despite the greater ease with which spoken language can now be investigated (see Section 1), it is probably fair to say that, as claimed by Linell (2005), linguistics is still very much characterized by a "written language bias": not only are most theories of language developed with written language in mind, but the discipline of linguistics relies more on the investigation of writing than speech. This can be explained by two main reasons (see also Gilquin and De Cock 2011). The first one is that capturing speech for linguistic analysis is generally more complex, more time-consuming and more expensive than capturing writing (think of the laborious task of transcribing speech verbatim, for example). Attempts at collecting spoken data have therefore been less common, and publicly available databases of speech are less numerous. The second reason is that writing has often been considered as the most prestigious form of language, more worthy of linguistic investigation than speech, which has frequently been relegated to the status of "stepchild of descriptive linguistics" (Tottie 1991: 255).[1]

Within the framework of Construction Grammar, too, writing tends to take center stage, to the detriment of speech. This can be illustrated by means of examples (1) to (4), quoted in the introductory chapter of Goldberg (2006: 6) and all taken from written sources. Such examples have served to develop the model of Construction Grammar, by demonstrating that abstract constructions carry meaning – one of the essential theoretical tenets of Construction Grammar. Although the constructions they represent are not limited to writing, such creative uses of the constructions are much more likely to occur in writing than in speech.

(1) He sneezed his tooth right across town. (Robert Munsch, *Andrew's Loose Tooth*, Scholastic Canada Ltd., 2002)
(2) She smiled herself an upgrade. (Douglas Adams, *Hitchhiker's Guide to the Galaxy*, New York: Harmony Books, 1979)
(3) We laughed our conversation to an end. (J. Hart, *Sin*, New York: Ivy Books, 1992)
(4) They could easily co-pay a family to death. (*New York Times*, January 14, 2002)

Two of the constructions that were investigated very early on in Construction Grammar, the *let alone* construction (cf. Fillmore, Kay, and O'Connor 1988) and the *way*-construction (cf. Israel 1996), are also more characteristic of the written register. While the former has a relative frequency of 13.84 per million words (pmw) in the written section of the British National Corpus and 5.67 pmw in the spoken section, the latter (in the form "V POSS. PR. *way* PREP.") has a relative frequency of 46.57 pmw in writing and 12.01 pmw in speech.[2]

[1] Paradoxically, however, speech can be argued to be the primary form of language: it developed long before the first writing systems appeared, it is acquired by children before they learn to write, language changes tend to occur in speech first, etc. (see Algeo and Pyles 2010: 6).

[2] As pointed out by a reviewer, the figures for speech might actually be slightly higher, because of possible intervening disfluencies that might prevent the retrieval of the pattern. A search for *er* and *erm* between the different elements of the constructions, however, did not produce any hits at all.

Several linguists have underlined the written bias in Construction Grammar. Linell (2005: 215) specifically mentions Construction Grammar when describing the written language bias in linguistics, noting that it "has so far not heeded the interactional aspects sufficiently". Fried and Östman (2005: 1753) refer to the "bias away from spoken language". More recently, Ginzburg and Poesio (2016: 1–2) claim the following, explicitly including Construction Grammar among contemporary linguistic theories:

> in practice contemporary theoretical linguistics is typically not interested or able to provide analyses for the rules governing language as it occurs in actual spoken interaction. Its analyses are developed for a cleaned up version of language […], which omits the disfluencies, interjections, overlapping turns, non-sentential utterances, and *ad hoc* coinages which are ubiquitous in spoken language […].

Yet, as a theoretical framework, Construction Grammar is perfectly suited to accommodate spoken phenomena. As Fried and Östman (2005: 1753) emphasize, there is no "built-in limitation of Construction Grammar with respect to its applicability to the study of spoken data and to larger stretches of discourse as the domain of analysis". In fact, speech has been involved in Construction Grammar right from the start. For example, in Fillmore, Kay, and O'Connor's (1988) paper, attention is paid to the typical prosody associated with the *let alone* construction, as in *She doesn't get up for LUNCH, let alone BREAKFAST*, where the capital letters represent prosodic peaks. As shown in Goldberg (2006: 14), other studies have focused on constructions that are more typical of speech, like the 'What's X doing Y' construction (e.g. *What's that fly doing in my soup?!*) investigated in Kay and Fillmore (1999), the 'Nominal Extraposition' construction (e.g. *It's amazing the difference!*) studied by Michaelis and Lambrecht (1996), or the so-called 'Mad Magazine' construction (e.g. *Him, a doctor?!*) explored in Lambrecht (1990). There is also a small number of Construction Grammar studies dealing with distinctively spoken phenomena, such as Fried and Östman (2005), Fischer (2010) and Fischer and Alm (2013) on pragmatic markers.

As for studies on child language acquisition (e.g. Diessel and Tomasello 2001), they are, for obvious reasons, always based on spoken data. In addition, most constructions, like the transitive, ditransitive, or intransitive motion constructions, are common in both speech and writing, and many corpus-based Construction Grammar studies "are based on a mixture of spoken and written data that characterizes most contemporary corpora" (Gries 2013: 97). Speech thus has a place in Construction Grammar, but in comparison with writing, there are fewer studies in Construction Grammar that deal specifically with speech and with typically spoken features.

If we now turn to Applied Construction Grammar and its focus on learner language, we find the same "bias away from spoken language" (Fried and Östman 2005: 1753), with most studies being based on experimental data and/or written corpus data (see Gilquin 2018: 128). The few studies that have relied on spoken data tend to be characterized by a small set of data, representing only a few learners, usually less than ten, and sometimes even just a single one, cf. Ellis and Ferreira-Junior (2009) on English verb-argument constructions, Eskildsen (2012) on English negation constructions, Eskildsen (2014) on the emergence of utterance schemas and schematic constructions, or Roehr-Brackin (2014) on German Perfekt constructions with *gehen* and *fahren*. Sambre et al. (2018) rely on a corpus made up of spoken data from 20 learners to investigate constructions expressing futurity. While Gilquin (2015) looks at more learners (almost 350) for a study of phrasal verb constructions, it is not exclusively concerned with speech, but with a combination of speech and writing. Naturally, as a relatively recent field, with a limited amount of research carried out so far, Applied Construction Grammar

cannot be blamed for not having explored all the aspects of language, including its different registers, but the above shows that it could do better justice to spoken phenomena. In what follows, we will consider several ways in which the study of speech can be beneficial to Applied Construction Grammar.

## 3 Descriptive perspective

This section examines how Applied Construction Grammar can benefit from the study of speech in terms of the description of constructions and of the constructicon. More precisely, it claims that we cannot dispense with the study of constructions in learner speech (Section 3.1) and that a comprehensive repertoire of spoken learner constructions (including disfluent and phonological constructions) is yet to be built (Section 3.2).

### 3.1 Comparison of spoken and written constructions

The stylistic sensitivity of constructions is a relatively well established fact by now. Constructions in native language tend to vary according to register (cf. Nikiforidou and Fischer 2015). This variation can affect different aspects of the construction, such as its frequency or its behavior. Smith (2009), for example, has demonstrated that light verb constructions with *have* and *make* are generally more frequent in speech than in writing. Shank, Plevoets, and Van Bogaert (2016), who study constructions with an object clause, note that *understand* is mostly used with a *that* complementizer in writing, but no complementizer in speech. Constructions in writing are thus not necessarily the same as constructions in speech.

Because learners have been claimed to lack stylistic awareness (Altenberg and Tapper 1998, Gilquin and Paquot 2008) and to use spoken-like features in writing and written-like features in speech, we can expect them to use constructions inappropriately in terms of register. This can be illustrated by means of the phrasal verb construction with *up* ([V *up*]) in speech and writing. Figure 1 shows the relative frequency per 100,000 words of the construction in the Louvain International Database of Spoken English Interlanguage (learner speech), the International Corpus of Learner English (learner writing), the Louvain Corpus of Native English Conversation (native speech), and the Louvain Corpus of Native English Essays (native writing). It appears from the graph that learners of English use the [V *up*] construction more often in writing than in speech, which is the opposite of the situation in native English, where the [V *up*] construction is more frequent in speech than in writing. This example shows that constructions may be used differently in native and learner speech.

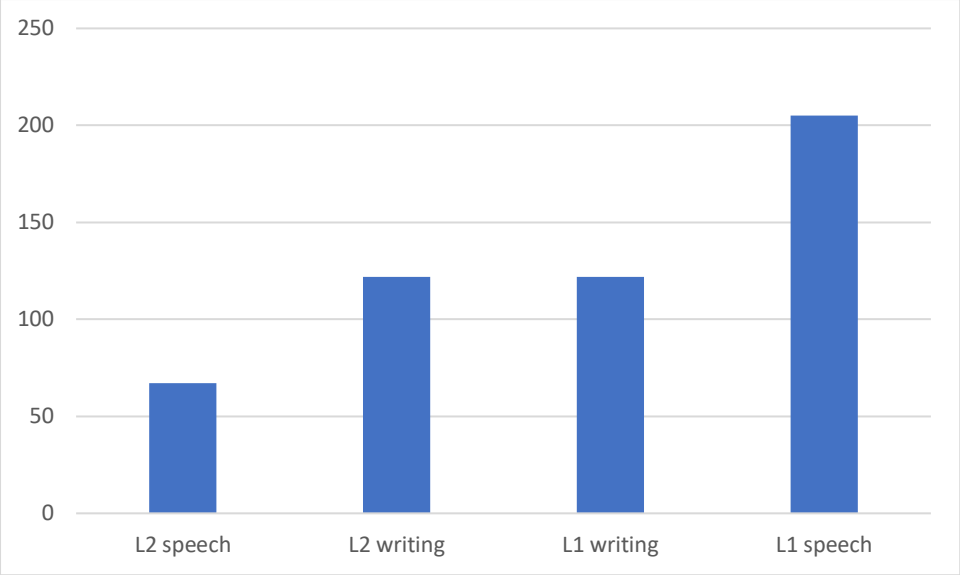

Figure 1. Relative frequency per 100,000 words of the [V *up*] construction in native (L1) and non-native (L2) speech and writing

If constructions in native writing may differ from constructions in native speech, and if constructions in native speech may differ from constructions in learner speech, then no shortcut from one to the other is possible. We can therefore not dispense with the study and the description of constructions in speech and in learner speech in particular. More generally, it is important to be aware of the possible impact of register on constructions and to bear in mind that constructions in speech may have their own specificities. The next obvious step, therefore, is to get to know more about the spoken learner constructicon: what does it look like? what constructions is it made up of? how can these constructions be described? We address this issue in the next section.

### 3.2 The spoken learner constructicon

While several projects have been launched to describe the constructicon of different languages (cf. Lyngfelt et al. 2018), as far as I know, no similar efforts have been made for learner language, let alone learner speech. We do have information about the use of linguistic phenomena in learner speech, such as clausal complementation in Tizón-Couto (2014), epistemic adverbial markers in Gablasova and Brezina (2015), or formulaic language in Wood (2010), all of which qualify as constructions in the Construction Grammar sense (although these studies are not theoretically embedded within Construction Grammar). However, these are isolated findings, which do not make it possible to approach learner speech from a global perspective. The spoken learner constructicon is thus still largely unexplored in Applied Construction Grammar.

In an attempt to start exploring the spoken learner constructicon, Gilquin (2018) carried out a study inspired by Cappelle and Grabar (2016), who suggested using recurrent sequences of part-of-speech (POS) tags, or POS n-grams, as an approximation to constructions, on the grounds that "patterns are stored as constructions even if they are fully predictable as long as they occur with sufficient frequency" (Goldberg 2006: 5). Cappelle and Grabar (2016) applied this methodology to native English, extracting the most frequent sequences of five POS tags from the Corpus of Contemporary American English. In Gilquin (2018), the methodology was applied to a part-of-speech tagged version of the Louvain International Database of Spoken English Interlanguage (LINDSEI; Gilquin, De Cock, and Granger 2010), representing 11

mother tongue populations, 554 different learners, and almost 800,000 words of L2 English spoken production. Recurrent sequences of 2 to 10 part-of-speech tags were extracted from the corpus taken as a whole (i.e. without making any distinction between the different L1 populations).

The results suggest that learners' spoken constructicon relies, in the first place, on short and simple constructions (mainly bigrams) of the type [NP], [Subj V] and [PP], and that the internal structure of these constructions seems relatively basic. The most frequent POS n-gram in LINDSEI corresponds to an NP which is made up of a determiner and a noun (e.g. *my brother*); the second and third POS n-grams correspond to VPs in which the subject consists in a personal pronoun, and the verb is either a lexical verb (e.g. *he left*) or the verb *be* (e.g. *I was*). More complex constructions of the type most often discussed in Construction Grammar, like argument structure constructions, are only found much further down the list of POS n-grams, which indicates that they are comparatively not very common in learner speech.[3] The spoken learner constructicon also turns out to be characterized by a great deal of disfluency. This appears from the presence of tags for interjections (predominantly corresponding to filled pauses) and truncations (i.e. unfinished words), as well as from incomplete sequences of tags, some of which are due to the spontaneous and unrehearsed nature of speech (e.g. ***in the*** *as we can see in the fourth picture*).

While the analysis in Gilquin (2018) was based on written transcriptions of the spoken data and could therefore not take phonological aspects into account, it stands to reason that a comprehensive account of the spoken learner constructicon should also include what could be referred to as phonological constructions. In this respect, I would like to mention a study by Drummond (2011), which is not embedded within the framework of Construction Grammar and does not use the term construction, but which is perfectly compatible with the tenets of Construction Grammar. The study considers glottal replacement (i.e. the replacement of /t/ by a glottal stop, /ʔ/) among Polish learners of English. Drummond identifies what he calls different 'environments' for potential glottal replacement, for example word final /t/ preceded by a vowel and followed by a word starting with a stop consonant (V/t/#S) as in 'h**o**t **p**otato' or word medial intervocalic /t/ as in 'ma**tt**er' (V/t/V). He then highlights some preferences among the learners in terms of glottal replacement in these different environments, such as the fact that it is almost inexistent in word medial position (V/t/V) or the fact that it represents only 17% in the V/t/#V (vowel + /t/ + word starting with a vowel, e.g. 'n**o**t **e**ven') and V/t/#P (vowel + /t/ + pause, e.g. 'did i**t**.') environments, to be compared with 40% and 36% respectively in native English. He also suggests that there might be a certain order of acquisition among learners, who seem to first use glottal replacement in the V/t/#V ('n**o**t **e**ven') environment, and only then in the V/t/#P ('did i**t**.') environment. The different environments identified by Drummond could be said to resemble constructions in the Construction Grammar sense, with a sequence of phones (vowel, stop consonant, etc.) corresponding to a certain meaning. The sequences with a /t/ vs /ʔ/ pronunciation are not unlike allostructions (Cappelle 2006), being two variants of a construction that share closely related semantics but produce different effects (e.g. in terms of formality). As for the preferences highlighted by Drummond, they could be systematized by means of a collostructional analysis, measuring the degree of attraction between phones (in this case, /t/ or /ʔ/) and a certain slot in the different constructions (or 'environments') identified.

[3] The results in Gilquin (2018) include a comparison with the POS n-grams in the native counterpart of LINDSEI, the Louvain Corpus of Native English Conversation (LOCNEC; De Cock 2004). It appears from the comparison that the spoken learner and native constructicons overlap to a certain extent, with both of them relying more on relatively basic constructions than on complex argument structure constructions.

Approaches such as those illustrated above can help build a map of the spoken learner constructicon, including disfluent constructions and phonological constructions. They also provide a useful basis to draw comparisons between spoken and written learner constructicons, as well as between spoken learner and native constructicons. The information thus collected makes it possible to fulfill a descriptive objective. In the next section, we consider the possible applications, among foreign language learners, of a Construction Grammar approach to speech.

## 4 Applied perspective

Studying learner speech from a Construction Grammar perspective can lead to useful applications. This section deals with the more cognitive/psycholinguistic aspects of foreign language learning (Section 4.1), as well as the pedagogical aspects of foreign language teaching (Section 4.2).

### 4.1 Understanding foreign language learning

Speech can help us understand the process of foreign language learning in ways that writing cannot, by providing better insights into the cognitive/psycholinguistic mechanisms that are at work during language production. The main reason for this is that written texts can be edited and redrafted a number of times, and that these changes are normally hidden from the final product. By contrast, speech is typically produced online, with no time to plan ahead. As a result, it may – and often does – include signs such as interruptions or restarts which show some of the processes that the speaker is going through when producing an utterance. Admittedly, even speech provides only a very partial window onto the psycholinguistic processes that characterize the production of language, as many things go on in the speaker's mind without being materialized by any actual language, but speech arguably gives a better understanding of these processes than writing.

The study of (spontaneous) learner speech can provide information about learners' degree of automaticity, a concept that is well-known in second language acquisition research and that describes the extent to which the application of rules has become automatic (cf. Segalowitz and Hulstijn 2005). Within the framework of Construction Grammar, automaticity can be said to correspond to the degree of entrenchment of a construction. If a construction is well entrenched, then the learner should be able to produce it automatically, in a rather smooth manner.

Examples (5) to (7) illustrate a gradual decrease in the degree of automaticity with which the learner applies the rule of the third person singular verb inflection. In (5), the learner seems to have internalized the rule and to apply it in a fully automatic manner; the construction is apparently well entrenched. In (6), the learner is able to correct herself thanks to attentional control, as indicated by the filled pause (*er*) followed by a restart; the application of the rule is not fully automatic and the construction is less entrenched. In (7), finally, the learner is either unaware of the rule or would need much more attentional control to be able to apply it; the entrenchment of the construction is non-existent or very limited.

(5) he's very successful cos **he works** hard in school (LINDSEI-CH)
(6) **he work** in a company (er) **he works** in a company (LINDSEI-CH)
(7) **he work** for the company for almost all his mature life (LINDSEI-CH)

Speech can also help identify constructions better, based for example on linguistic elements that are suggestive of language chunking. Thus, filled pauses have been said to be

"indirect indicators of prefabricated language and holistic storage" (Dahlmann and Adolphs 2009: 126) – a description that is reminiscent of Goldberg's (2003: 219) definition of constructions as "*stored* pairings of form and function" (emphasis added). Filled pauses can therefore inform us about the constructions that are likely to be stored – or not – in the learner's mind.[4]

The POS n-gram-based exploration of the spoken learner constructicon in Gilquin (2018) revealed that filled pauses sometimes interrupt closely-knit structures in English as a foreign language. Table 1, based on the analysis of the Louvain Corpus of Native English Conversation (LOCNEC; De Cock 2004) for native speech (NS) and of LINDSEI for non-native speech (NNS), shows the distribution of a number of patterns centered around a prepositional phrase made up of a preposition (Prep), a determiner (Det) and a noun (N), and interrupted or not by a filled pause (UH). In a majority of cases in both native and non-native speech, the sequence is not interrupted by any pause and is not preceded nor followed by any pause either, which suggests that for most native and non-native speakers the prepositional phrase is a well-entrenched construction. The proportion of such cases, however, is much higher in NS (94.8%) than in NNS (79.2%). This is to be related to the fact that in NNS the prepositional phrase is regularly preceded, followed or interrupted by a pause.

Interruptions of the prepositional phrase by a pause are particularly interesting, because they suggest that among certain speakers, the PREP-DET-N sequence may not be stored as a construction, or at least may not be well entrenched in the mind. In NNS the interruption between the preposition and the determiner, as in *to **er** the lady*, accounts for more than 3%, to be compared with 0.6% in NS. The proportion of interruptions between the determiner and the noun, as in *to our **erm** bus*, is lower, but still higher in NNS (1.7%) than in NS (0.3%). This difference between the two interrupted patterns, incidentally, also suggests that there are fewer learners for whom the noun phrase may not be stored as a construction, since the determiner and the noun are less often interrupted by means of a pause than the preposition and the noun phrase.

The same sort of phenomenon can be found with discourse markers, which also sometimes interrupt closely-knit structures and can be indicative of the (lack of) storage of certain constructions. Gilquin and Granger (2015: 434) found out that the discourse marker *you know* was overused by Polish learners of English in LINDSEI in comparison to native speakers in LOCNEC, and that the proportion of interrupted uses was much higher among the learners (60%) than among the native speakers (35%). The learners, for instance, used the discourse marker between a copula and the predicative (e.g. *he stopped being **you know** humorous*; *it was always **you know** a kind of foreigner*) or between a preposition and the rest of the prepositional phrase (e.g. *I changed my mind under the pressure of of **you know** society*; *that was some guy from **you know** the upper classes*). Such interruptions seem to indicate that constructions of a higher level like prepositional phrases may not be stored among some learners, and that their entrenchment is weaker than that of lower-level constructions like individual words.

| **Structure** | | | | | | | **Examples** | **NS** | **NNS** |
|---|---|---|---|---|---|---|---|---|---|
| | Prep | | Det | | N | | *about a book* | 94.8% | 79.2% |
| UH | Prep | | Det | | N | | ***eh** in the film* | 1.9% | 8.6% |
| | Prep | UH | Det | | N | | *to **er** the lady* | 0.6% | 3.2% |
| | Prep | | Det | UH | N | | *to our **erm** bus* | 0.3% | 1.7% |
| | Prep | | Det | | N | UH | *in a hotel **er*** | 2.4% | 7.3% |

[4] Note that pauses can also be symptomatic of other problems, such as lexical retrieval difficulties, and that in some cases pauses may be intentional (e.g. to create rhetorical emphasis).

Table 1. Percentage of interrupted and uninterrupted uses of the prepositional phrase [Prep Det N] in LOCNEC (NS) and LINDSEI (NNS)

The fact that foreign language learners may rely more on lower-level constructions than on higher-level constructions, and in particular more on words than on phrases, ties in with Wible's (2008: 166–167) claim about multiword expressions. He argues that learners who depend mainly on spoken input tend to go through a "whole to parts" process, which involves the segmentation of the speech stream into individual words. Learners who rely mainly on written input, on the other hand, are more likely to go through a "parts to whole" process, starting from individual words, clearly delineated from each other in writing, and building up multiword expressions on this basis. This may point to a difference in the way foreign language learners, on the one hand, and children and second language learners, on the other, build their constructicons.

In this view, foreign language learners, whose main input in the target language tends to be written (see Section 4.2), can be hypothesized to start from lower-level constructions (especially words) and to gradually build up higher-level constructions on this basis. Children and second language learners, by contrast, are primarily exposed to the target language in its spoken form, and can therefore be hypothesized to first acquire higher-level constructions, which they then have to segment into lower-level constructions. The so-called "emergent lexicon" (Bybee 1998: 434), according to which "[t]he smaller units […] emerge from these larger stored units via a network of connections among them", may thus apply to children acquiring their mother tongue and to second learners of a language, but perhaps not to foreign language learners, whose primary input in the target language is usually written.

The phonetic aspects of speech can also provide insights into the processing of constructions, and more particularly into the ease with which constructions are processed, depending on their probability in the (learner) constructicon. Thus, Kuperman and Bresnan (2012: 590) point out that "low-level continuous phonetic variation is sensitive to high-level construction probabilities". This has been demonstrated for both elicited and spontaneous speech. In elicited speech based on a visual cue, it appears that when subjects produce a variant that, on the basis of the linguistic environment, is less likely, the voice initiation time, i.e. the time lag between the visual cue and the onset of speech, is longer (Francis and Michaelis 2017). In spontaneous speech, the production of a less probable variant, given a certain linguistic context, tends to correspond to longer acoustic duration (Kuperman and Bresnan 2012). Reduction effects in speech point in the same direction, since "words are pronounced in a more reduced fashion if they occur in a construction of which they are highly typical" (Hilpert 2014: 144); words that occur in a construction of which they are not typical, conversely, are less likely to display such reduction effects.

Longer voice initiation time in elicited speech and longer acoustic duration in spontaneous speech are indicative of processing effort, which itself depends on the probability (or rather, improbability) of the construction. Although these studies concern native speech, it seems reasonable to assume that, in learner speech too, phonetic features such as acoustic duration might be clues about processing ease and construction probability for the learner. These indicators could be compared with those found in native speech, in an attempt to highlight the specificities of learner speech processing.

Finally, it is worth considering the speech that learners get exposed to. Examining the spoken form of language input can help understand the acquisition process, by explaining why certain constructions may be easier to learn in some contexts. Thus, it has been said that words following the filled pause *er* are easier to remember (Corley, MacGregor, and Donaldson 2007). Phonological similarity, in the form of alliteration patterns, is also claimed to be relevant for language learning (cf. Gries 2011: 504). Phonetic and perceptual salience seems to be useful

too, since it has been shown that constructions that are salient in learners' language input (for example by being stressed or more clearly articulated) tend to be acquired earlier (Collins et al. 2009). Such features may therefore help predict the ease with which constructions will be acquired or memorized.

This section has sought to demonstrate that speech offers numerous clues that can help understand the processes underlying foreign language learning: not only can the study of learner speech and its features (e.g. use of pauses or discourse markers, acoustic duration) tell us something about how learners process constructions or how their constructicon develops, but, in addition, the features of the spoken language input that learners receive can help predict their acquisition of constructions. Another interesting field of application, related to foreign language learning, is pedagogy. In the next section, we will see how the study of speech can help improve the teaching of constructions.

### 4.2 Improving foreign language teaching

Generally speaking, we can say that foreign language learners get relatively little exposure to (authentic) speech in the target language. In a typical foreign language situation, the target language is learned via instruction, and spoken input is often limited to teacher talk and oral pedagogical materials (Gilquin and Paquot 2008: 52). Some years ago, comparisons of textbooks and authentic speech emphasized the lack of authenticity of dialogues in textbooks (cf. Gilmore 2004, Römer 2004).

Things have changed slightly since then, both in terms of the quality of textbooks and additional opportunities that foreign language learners may have to receive spoken or spoken-like language input, especially for English (e.g. through TV series). However, it is still true that authentic speech does not usually represent the main component of the language input that foreign language learners receive. In addition, it was underlined in Section 3.1 that learners tend to be unaware of stylistic differences between constructions. It is therefore important that learners should get exposed to spoken constructions and should be made aware of their distinctive features and contexts of use. Typically spoken phenomena should also be incorporated into the foreign language curriculum (see also Gilquin and De Cock 2011: 159), including, for example, discourse markers, which are essential to "oil […] the wheels of verbal interaction" (Stubbe and Holmes 1995: 63), but which are often underused and/or misused by foreign language learners (see, e.g., Aijmer 2011 or Gilquin 2016).

Potentially relevant from a pedagogical point of view are the recent approaches of dialogic syntax (Du Bois 2014) and dialogic Construction Grammar (Brône and Zima 2014), which consider that constructions can span several turns, thus relying on the interactional nature of speech. If we combine these approaches with De Pietro, Matthey, and Py's (1989) concept of "séquences potentiellement acquisitionnelles" (*sequences of potential acquisition*; see Perdue 2000), which describes interactional sequences conducive to learning, we can propose a dialogic construction for corrective feedback. Such a construction is illustrated in (8), taken from the Spanish component of LINDSEI, where A represents the interviewer and B represents the learner.

(8)

| | |
|---|---|
| B: he's he's using . some .. I don't know how do you say this . these things you use to paint | Communicative breakdown |
| A: a **brush** | Expert intervention |
| B: **brush** okay yes thank you<br>(A: mhm)<br>B: yes she he's using some **brushes** to to paint the picture | Incorporation of feedback |

(LINDSEI-SP)

The first turn corresponds to a communicative breakdown: the learner does not know how to say 'brush' and uses a paraphrase to describe the object (*these things you use to paint*), while clearly signaling the problem (*I don't know how do you say this*). The second turn shows the expert intervention, where the interviewer provides the right word (*a brush*). The learner then incorporates the expert's feedback by repeating the word *brush* and acknowledging that she has understood (*okay yes*). In her next turn, she integrates the word into the sentence (*he's using some brushes to to paint the picture*). The same type of dialogic construction is found twice in example (9), taken from the French component of LINDSEI, where the learner uses her mother tongue (cf. words in italics) to get the expert to help her.

(9)

| | CONSTRUCTION 1 |
|---|---|
| B: I think they have er this room for for the whole family with er just erm .. erm <starts whispering> er *fauteuil* how do you say that | Communicative breakdown |
| A: a **chair** | Expert intervention |
| B: a **chair** yes . some chairs for the the family in er . in the: in the corner and just erm .. er .. | Incorporation of feedback |
| | CONSTRUCTION 2 |
| B: *oh la la* how do you say that er | Communicative breakdown |
| A: a **bookcase** *une armoire* a cupboard | Expert intervention |
| B: yes a **bookcase** | Incorporation of feedback |

(LINDSEI-FR)

Interaction, including teacher-learner interaction, can facilitate foreign language learning (Verga and Kotz 2013). Dialogic constructions building on such interaction, for instance to provide corrective feedback, are regularly used by teachers and they should be encouraged in the classroom or in any other setting where learning can take place. While they have already been examined from other theoretical perspectives (e.g. Mackey, Oliver, and Leeman 2003), their formalization within Construction Grammar could be a first step toward their closer investigation by (applied) construction grammarians.

Taking pedagogical intervention one step further, the teacher can engage in instructional manipulation. As Collins et al. (2009: 348) point out, "one of the distinct advantages of the classroom is the possibility of facilitating acquisition by manipulating input and output opportunities for learners". One well-known type of manipulation is the repeated exposure of learners to the same construction, as recommended by Nation (2001) for vocabulary acquisition. Following up on Collins et al.'s (2009) study on phonetic and perceptual salience (see Section 4.1), we could also recommend teachers to increase the salience of opaque constructions, e.g. through stress, intonation, or articulation, in order to help students notice these constructions and remember them better. Filled pauses could be used with the same effect (cf. Corley, MacGregor, and Donaldson 2007, Section 4.1). Presenting learners with written transcripts of spoken texts might be another useful type of manipulation, not suggested by Collins et al. (2009) but building on their observation that "encounters with new features in spoken contexts are fleeting, offering listeners little opportunity to reflect on the linguistic properties of what they hear" (Collins et al. 2009: 350). A written transcript could make these fleeting encounters more lasting, thus giving learners time to consider constructions that are typical of spoken language.

Finally, the study of spoken data could change attitudes towards non-standard language. As pointed out by Lintunen and Mäkilä (2014: 378), "higher expectations are set for accuracy in writing, whereas spoken language allows many inaccuracies in form". By examining spoken

native data, teachers and/or researchers might become more tolerant toward certain aspects of learner language, including what Waara (2004) calls "learner constructions", that is, utterances that are "slightly off" and consist in "form-meaning reconstructions" (Waara 2004: 51). An example would be the [*DISCUSS about* NP] construction, illustrated in (10), which is taken from the Finnish component of the International Corpus of Learner English. While such a construction is normally considered an error that should be corrected in learner language, the fact that it can be found in spoken native data (cf. (11), taken from the Corpus of Contemporary American English) might make some teachers hesitant to reject it categorically among their learners.

(10) Although the jury members are not allowed to follow the course of the events in the media or to **discuss about** the trial with each other, there are no garantees about that they will not do that. (ICLE-FI)
(11) Well, you come next year and then we'll **discuss about** that (COCA-SPEECH)

Similarly, Götz (2013), in a study on the fluency of native and non-native English speech, has shown that, although on average native speakers of English use fewer filled pauses than learners of English, there is some overlap between the two groups, with some native speakers actually using more filled pauses, and thus being arguably more disfluent, than certain learners. This seems to be the case if we compare example (12), uttered by a German learner, and example (13), uttered by a native speaker. While the former only includes one filled pause (and a vowel lengthening on *the*), the latter includes as many as three filled pauses, which are combined with other disfluent features (unfilled pause [indicated by the dot], repeat, vowel lengthening). Again, such findings could make teachers more tolerant toward the imperfections of their learners' language production.[5]

(12) I chose the movie Goodbye Lenin which is a movie about the[i:] **erm** German reunification (LINDSEI-GE)
(13) the film is basically about . **erm** the life of William Wallis who was a: a leader **erm** he was **erm** a fairly simple person (LOCNEC)

In this section, we have seen that the study of speech in Applied Construction Grammar can bring pedagogical benefits. It can encourage teachers to focus more on spoken constructions and stylistic issues. The manipulation of learners' language input (repeated exposure to certain spoken constructions, emphasis through articulation or intonation, etc.) or the use of dialogic constructions (e.g. for corrective feedback) can also enhance the comprehension and acquisition of constructions. As for the study of native speech, it can increase our tolerance toward non-standard features in learner language. Having considered the applied perspective, we now move on to the final dimension, that of the possible theoretical implications of a better recognition of speech in Applied Construction Grammar.

## 5 Theoretical perspective

Including more speech in Applied Construction Grammar can have an impact on the model of Construction Grammar itself. One change that the study of spoken data can bring about is a

[5] This is not to say that teachers should necessarily tolerate in their learners' production anything that can be found in native speech. Maintaining certain standards in the foreign language classroom is arguably desirable in many teaching contexts.

shift to the more usual aspects of language, which reflects Bergs's (2008) question: "Can we take Construction Grammar beyond sneezing napkins off tables?". Goldberg (2003: 291) points out that "unusual constructions shed light on more general issues, and can illuminate what is required for a complete account of language". However, we should also take the other side of the coin into account and recognize, as Stubbs (2002: 221) does, that "[u]nique events certainly occur, but can be described only against the background of what is normal and expected".

The POS n-gram analysis of LINDSEI described in Section 3.2 brought to light constructions that are so simple that they may almost seem disappointing in comparison with the types of constructions that are typically highlighted in Construction Grammar. This does not appear to be limited to learner speech, as native speech (at least native students' speech) also seems to rely predominantly on rather basic constructions (Gilquin 2018). Yet, it is important to bear in mind that the more mundane aspects of language have an essential role to play and thus deserve to take center stage in linguistic analysis.

Taking speech into account will also allow for a welcome acknowledgment of the phonological aspects of constructions. Phonological properties are part and parcel of the symbolic structure of constructions, as appears from Figure 2, taken from Croft (2001: 18), as well as Croft and Cruse's (2004: 247) statement that "[c]onstructions, like the lexical items in the lexicon, are 'vertical' structures that combine syntactic, semantic and even phonological information (for the specific words in a construction, as well as any unique prosodic features that may be associated with a construction)". However, as pointed out by Boas (2013: 239), "[w]hile constructional research has focused primarily on the role of semantic, pragmatic, and syntactic factors in licensing constructions, very few studies address the influence of phonological factors". Boas (2013: 240) adds that "phonological information may sometimes also be relevant for the licensing of constructions, and as such need to be included in formal representations when necessary". This can only be done through the careful analysis of spoken corpora, or more precisely speech corpora, which give access to the acoustic signal in addition to the transcripts.

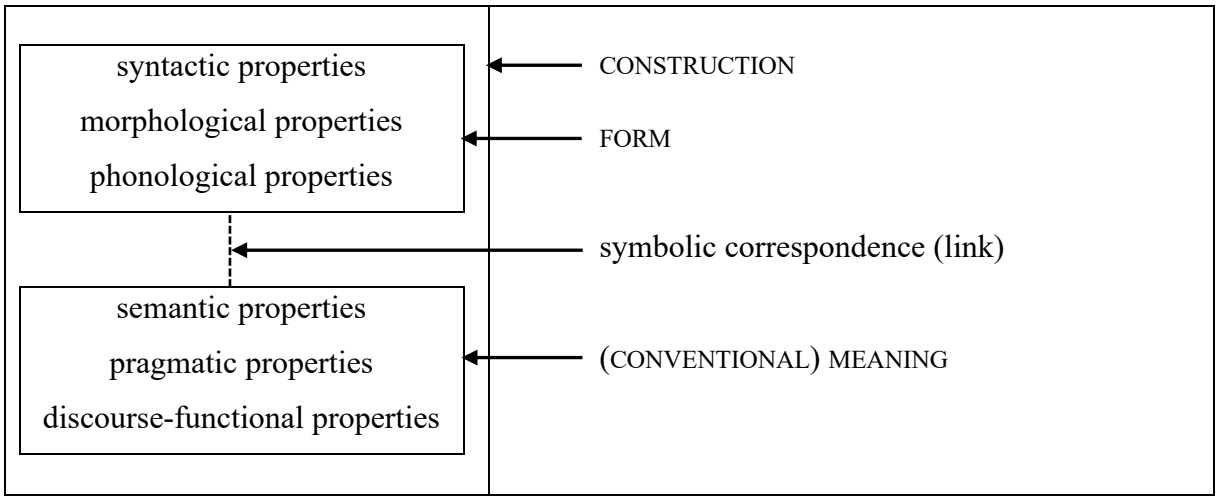

Figure 2. Representation of the symbolic structure of a construction (Croft 2001: 18)

Figure 2 also shows that stylistic features are part of the meaning pole of constructions, in the form of discourse-functional properties. Just like phonological properties, however, stylistic aspects, including differences between speech and writing, have often been neglected in Construction Grammar. Bergs and Diewald (2009a: 2) pointed out that "[s]o far, many studies in construction grammar have focused on (verb) semantic aspects. […] most [studies] have treated pragmatics and *discourse* more or less in passing, so that there is still no systematic survey of this domain" (emphasis added). Their book (Bergs and Diewald 2009b) was a first

attempt to consider these aspects more systematically in Construction Grammar. Recently, there has also been a growing "interest in genre-based or genre-sensitive constructions" (Nikiforidou and Fischer 2015: 139). More research on speech in Construction Grammar should further stimulate this interest in stylistic issues.

Research on speech can also help uncover new types of constructions. Phonological constructions were discussed in Section 3.2 and dialogic constructions in Section 4.2. Pragmatic markers have also been claimed to correspond to constructions, most notably by Fried and Östman (2005) and Fischer (2010). Filled pauses, which have been mentioned several times in what precedes, could be said to function as constructions too. In the literature, they have been described as *bona fide* words (Tottie 2015) – and since words are constructions, filled pauses must be constructions. Like constructions, they consist in a pairing of form (e.g. *er* or *erm*) and function (for example, keep the floor, request help to complete the utterance, invite the addressee to speak, or show politeness; see Clark and Fox Tree 2002 for a list of some of the functions of filled pauses). Unfilled pauses are more difficult to deal with. While they undoubtedly fulfill certain functions (cf. Esposito et al. 2007), some might object to silence being considered a linguistic form. And yet, if "silence speaks volumes", it must be a form of communication.[6]

At a more general level, the use of spoken data can help get closer to the ideal of psychological plausibility that characterizes (the more cognitive strand of) Construction Grammar (Boas 2013: 248). While experimentation is arguably the most suitable method to achieve this aim, among corpus data, speech may be said to be a better indicator of psycholinguistic processes, such as construction-forming or automaticity, than writing (see also Section 4.1). Speech can thus promote a more psychologically plausible form of (Applied) Construction Grammar.

This section has explored the idea that speech can contribute to (Applied) Construction Grammar as a theoretical model. Its study can help refine the model so that it explicitly accommodates the specificities of (learner) speech. It also draws our attention to properties of constructions that tend to be neglected, especially phonological and discourse-functional properties, and can lead to the discovery of new types of constructions. Finally, it can bring us closer to the ideal of psychological plausibility that (Cognitive) Construction Grammar strives for.

## 6 Conclusion

Lintunen and Mäkilä (2014: 377) rightly point out that "[t]here is an axiomatic difference between speech and writing, but together they form the essential parts of learners' L2 skills". Speech, therefore, deserves the attention of (applied) linguists and teaching practitioners just as much as writing. Adopting the perspective of Applied Construction Grammar, this paper has shown that research on speech can help provide better descriptions of spoken learner constructions, which may differ from written learner constructions or spoken native constructions, and can help approach the spoken learner constructicon. Spoken language research can also offer insights into the cognitive mechanisms underlying L2 acquisition, answering questions such as "how do learners build their constructicon?" or "how do they process constructions?". In addition, it can contribute to the improvement of teaching practices by bringing stylistic issues to the forefront or by manipulating learners' language input. Finally,

[6] During the discussion time that followed my plenary talk at CALP3, on which this paper is based, Thomas Herbst asked the interesting question of whether accent could be considered a construction. The jury is still out on whether this is the case or not.

it can help refine the model of (Applied) Construction Grammar and widen its scope, by explicitly making room for typically spoken phenomena like interaction or disfluencies.

If research on speech can prove to be useful on all of these accounts, it is well worth devoting more time and energy to it in the developing field of Applied Construction Grammar. Of course, speech cannot be studied in isolation from writing. Research on written language therefore needs to be pursued in Construction Grammar, so that the distinctive features of speech can be easily highlighted. It is also important to take into account that speech and writing represent a continuum rather than a dichotomy, and that many of the new emerging genres, like tweets (cf. Paveau 2013) or TED talks (cf. Ludewig 2017), are mixed genres in terms of the mode they represent, combining features of speech and writing. This grey area between the spoken and written poles, as well as the evolution of modes, will have to enter the equation too, and should be on the research agenda of (applied) construction grammarians. Clearly, there's plenty to do, so let's get down to it!

**References**


Aijmer, Karin. 2011. *Well I'm not sure I think…* The use of *well* by non-native speakers. *International Journal of Corpus Linguistics* 16 (2). 231–254.

Algeo, John and Thomas Pyles. 2010. *The origins and development of the English language*. Sixth Edition. Boston, MA: Wadsworth.

Altenberg, Bengt & Marie Tapper. 1998. The use of adverbial connectors in advanced Swedish learners' written English. In Sylviane Granger (ed.), *Learner English on computer*, 80–93. London & New York: Addison Wesley Longman.

Bergs, Alexander. 2008. Can we take Construction Grammar beyond sneezing napkins off tables? In Klaus Stierstorfer (ed.), *Proceedings of the Anglistentag Münster 2007*, 269–276. Trier: WVT.

Bergs, Alexander & Gabriele Diewald. 2009a. Contexts and constructions. In Alexander Bergs & Gabriele Diewald (eds.), *Contexts and constructions*, 1–14. Amsterdam & Philadelphia: John Benjamins.

Bergs, Alexander & Gabriele Diewald (eds.). 2009b. *Contexts and constructions*. Amsterdam & Philadelphia: John Benjamins.

Boas, Hans C. 2013. Cognitive Construction Grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Grammar*, 233–252. Oxford: Oxford University Press.

Brône, Geert & Elisabeth Zima. 2014. Towards a dialogic construction grammar: *Ad hoc* routines and resonance activation. *Cognitive Linguistics* 25 (3). 457–495.

Bybee, Joan. 1998. The emergent lexicon. *Chicago Linguistic Society* 34. 421–435.

Cappelle, Bert. 2006. Particle placement and the case for "allostructions". *Constructions* SV1-7/2006.

Cappelle, Bert & Natalia Grabar. 2016. Towards an n-grammar of English. In Sabine De Knop & Gaëtanelle Gilquin (eds.), *Applied Construction Grammar*, 271–302. Berlin: De Gruyter.

Clark, Herbert H. & Jean E. Fox Tree. 2002. Using *uh* and *um* in spontaneous speaking. *Cognition* 84 (1). 73–111.

Collins, Laura, Pavel Trofimovich, Joanna White, Walcir Cardoso & Marlise Horst. 2009. Some input on the easy/difficult grammar question: An empirical study. *The Modern Language Journal* 93 (3). 336–353.

Corley, Martin, Lucy J. MacGregor & David I. Donaldson. 2007. It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition* 105 (3). 658–668.

Croft, William. 2001. *Radical Construction Grammar. Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Croft, William & D. Alan Cruse. 2004. *Cognitive linguistics*. Cambridge: Cambridge University Press.
Dahlmann, Irina & Svenja Adolphs. 2009. Spoken corpus analysis: Multimodal approaches to language description. In Paul Baker (ed.), *Contemporary corpus linguistics*, 125–139. London & New York: Continuum.
De Cock, Sylvie. 2004. Preferred sequences of words in NS and NNS speech. *Belgian Journal of English Language and Literatures (BELL), New Series* 2. 225–246.
De Knop, Sabine & Gaëtanelle Gilquin (eds.). 2016. *Applied Construction Grammar*. Berlin: De Gruyter.
de Pietro, Jean-François, Marinette Matthey & Bernard Py. 1989. Acquisition et contrat didactique : les séquences potentiellement acquisitionnelles de la conversation exolingue. In Dominique Weil & Huguette Fugier (eds.), *Actes du troisième colloque régional de linguistique, Strasbourg, 28–29 avril 1988*, 99–124. Université des Sciences Humaines/Université Louis Pasteur.
Diessel, Holger & Michael Tomasello. 2001. The acquisition of finite complement clauses in English: A corpus-based analysis. *Cognitive Linguistics* 12 (2). 97–141.
Drummond, Rob. 2011. Glottal variation in /t/ in non-native English speech: Patterns of acquisition. *English World-Wide* 32 (3). 280–308.
Du Bois, John W. 2014. Towards a dialogic syntax. *Cognitive Linguistics* 25 (3). 359–410.
Ellis, Nick C. & Fernando Ferreira-Junior. 2009. Constructions and their acquisition. Islands and the distinctiveness of their occupancy. *Annual review of Cognitive Linguistics* 7. 187–220.
Eskildsen, Søren Wild. 2012. L2 negation constructions at work. *Language Learning* 62 (2). 335–372.
Eskildsen, Søren Wild. 2014. What's new? A usage-based classroom study of linguistic routines and creativity in L2 learning. *International Review of Applied Linguistics* 52 (1). 1–30.
Esposito, Anna, Vojtěch Stejskal, Zdeněk Smékal & Nikolaos Bourbakis. 2007. The significance of empty speech pauses: Cognitive and algorithmic issues. In Francesco Mele, Giuliana Ramella, Silvia Santillo & Francesco Ventriglia (eds.), *Advances in brain, vision, and artificial intelligence. Proceedings of the second international symposium, BVAI 2007, Naples, Italy, October 2007*, 542–554. Berlin & Heidelberg: Springer-Verlag.
Fillmore, Charles J., Paul Kay & Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64 (3). 501–538.
Fischer, Kerstin. 2010. Beyond the sentence. Constructions, frames and spoken interaction. *Constructions and Frames* 2 (2). 185–207.
Fischer, Kerstin & Maria Alm. 2013. A radical construction grammar perspective on the modal particle-discourse particle distinction. In Liesbeth Degand, Bert Cornillie & Paola Pietrandrea (eds.), *Discourse markers and modal particles: Categorization and description*, 47–87. Amsterdam & Philadelphia: John Benjamins.
Francis, Elaine J. & Laura A. Michaelis. 2017. When relative clause extraposition is the right choice, it's easier. *Language and Cognition* 9 (2). 332–370.
Fried, Mirjam & Jan-Ola Östman. 2005. Construction Grammar and spoken language: The case of pragmatic particles. *Journal of Pragmatics* 37 (11): 1752–1778.
Gablasova, Dana & Vaclav Brezina. 2015. Does speaker role affect the choice of epistemic adverbials in L2 speech? Evidence from the Trinity Lancaster Corpus. In Jesús Romero-Trillo (ed.), *Yearbook of corpus linguistics and pragmatics 2015: Current approaches to discourse and translation studies*, 117–136. Dordrecht: Springer.
Gilmore, Alex. 2004. A comparison of textbook and authentic interactions. *ELT Journal* 58 (4). 363–374.

Gilquin, Gaëtanelle. 2015. The use of phrasal verbs by French-speaking EFL learners. A constructional and collostructional corpus-based approach. *Corpus Linguistics and Linguistic Theory* 11 (1). 51–88.
Gilquin, Gaëtanelle. 2016. Discourse markers in L2 English: From classroom to naturalistic input. In Olga Timofeeva, Anne-Christine Gardner, Alpo Honkapohja & Sarah Chevalier (eds.), *New approaches to English linguistics: Building bridges*, 213–249. Amsterdam & Philadelphia: John Benjamins.
Gilquin, Gaëtanelle. 2018. Exploring the spoken learner English constructicon: A corpus-driven approach. In Rosa Alonso Alonso (ed.), *Speaking in a second language*, 127–152. Amsterdam & Philadelphia: John Benjamins.
Gilquin, Gaëtanelle & Sylvie De Cock. 2011. Errors and disfluencies in spoken corpora: Setting the scene. *International Journal of Corpus Linguistics* 16 (2). 141–172.
Gilquin, Gaëtanelle, Sylvie De Cock & Sylviane Granger. 2010. *Louvain International Database of Spoken English Interlanguage. Handbook and CD-ROM*. Louvain-la-Neuve: Presses universitaires de Louvain.
Gilquin, Gaëtanelle & Sabine De Knop. 2016. Exploring L2 constructionist approaches. In Sabine De Knop & Gaëtanelle Gilquin (eds.), *Applied Construction Grammar*, 3–17. Berlin: De Gruyter.
Gilquin, Gaëtanelle & Sylviane Granger. 2015. Learner language. In Douglas Biber & Randi Reppen (eds.), *The Cambridge handbook of English corpus linguistics*, 418–435. Cambridge: Cambridge University Press.
Gilquin, Gaëtanelle & Magali Paquot. 2008. Too chatty: Learner academic writing and register variation. *English Text Construction* 1 (1). 41–61.
Ginzburg, Jonathan & Massimo Poesio. 2016. Grammar is a system that characterizes talk in interaction. *Frontiers in Psychology* 7. 1938. doi:10.3389/fpsyg.2016.01938.
Goldberg, Adele E. 2003. Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences* 7 (5). 219–224.
Goldberg, Adele E. 2006. *Constructions at work. The nature of generalization in language*. Oxford: Oxford University Press.
Götz, Sandra. 2013. *Fluency in native and nonnative English speech*. Amsterdam & Philadelphia: John Benjamins.
Gries, Stefan Th. 2011. Phonological similarity in multi-word units. *Cognitive Linguistics* 22 (3). 491–510.
Gries, Stefan Th. 2013. Data in Construction Grammar. In Thomas Hoffmann & Graeme Trousdale (eds.), *The Oxford handbook of Construction Grammar*, 93–108. Oxford: Oxford University Press.
Hilpert, Martin. 2014. *Construction Grammar and its application to English*. Edinburgh: Edinburgh University Press.
Israel, Michael. 1996. The way constructions grow. In Adele E. Goldberg (ed.), *Conceptual structure, discourse and language*, 217–230. Stanford, CA: CSLI Publications.
Kay, Paul & Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: The *What's X doing Y?* construction. *Language* 75 (1). 1–33.
Kuperman, Victor & Joan Bresnan. 2012. The effects of construction probability on word durations during spontaneous incremental sentence production. *Journal of Memory and Language* 66 (4). 588–611.
Lambrecht, Knud. 1990. 'What, me worry?': Mad magazine sentences revisited. In Kira Hall, Jean-Pierre Koenig, Michael Meacham, Sondra Reinman & Laurel A. Sutton (eds.), *Proceedings of the sixteenth annual meeting of the Berkeley Linguistics Society*, 215–228. Berkeley: Berkeley Linguistics Society.

Linell, Per. 2005. *The written language bias in linguistics. Its nature, origins and transformations*. London & New York: Routledge.
Lintunen, Pekka & Mari Mäkilä. 2014. Measuring syntactic complexity in spoken and written learner language: Comparing the incomparable? *Research in Language* 12 (4). 377–399.
Ludewig, Julia. 2017. TED Talks as an emergent genre. *CLCWeb: Comparative Literature and Culture* 19 (1). http://docs.lib.purdue.edu/clcweb/vol19/iss1/2 (accessed 29 October 2018).
Lyngfelt, Benjamin, Lars Borin, Kyoko Ohara & Tiago Timponi Torrent (eds.). 2018. *Constructicography: Constructicon development across languages*. Amsterdam & Philadelphia: John Benjamins.
Mackey, Alison, Rhonda Oliver & Jennifer Leeman. 2003. Interactional input and the incorporation of feedback: An exploration of NS–NNS and NNS–NNS adult and child dyads. *Language Learning* 53 (1). 35–66.
Michaelis, Laura A. & Knud Lambrecht. 1996. Toward a construction-based theory of language function: The case of nominal extraposition. *Language* 72 (2). 215–247.
Nation, Paul. 2001. *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
Nikiforidou, Kiki & Kerstin Fischer. 2015. On the interaction of constructions with register and genre. *Constructions and Frames* 7 (2). 137–147.
Paveau, Marie-Anne. 2013. Genre de discours et technologie discursive. Tweet, twittécriture et twittérature. *Pratiques* 157-158. 7–30.
Perdue, Clive. 2000. Untutored language acquisition. In Michael Byram (ed.), *Routledge encyclopedia of language teaching and learning*, 939–947. London & New York: Routledge.
Roehr-Brackin, Karen. 2014. Explicit knowledge and processes from a usage-based perspective: The developmental trajectory of an instructed L2 learner. *Language Learning* 64 (4). 771–808.
Römer, Ute. 2004. Comparing real and ideal language learner input: The use of an EFL textbook corpus in corpus linguistics and language teaching. In Guy Aston, Silvia Bernardini & Dominic Stewart (eds.), *Corpora and language learners*, 151–168. Amsterdam & Philadelphia: John Benjamins.
Sambre, Paul, Julien Perrez, Pascale Van Keirsbilck & Cornelia Wermuth. 2018. Constructing futurity: A contrastive approach to L1 and L2 Dutch and French. Paper presented at *Constructionist Approaches to Language Pedagogy 3 (CALP 3)*, 15–17 February 2018, Austin, TX.
Segalowitz, Norman & Jan Hulstijn. 2005. Automaticity in bilingualism and second language learning. In Judith F. Kroll & Annette M. B. De Groot (eds.), *Handbook of bilingualism. Psycholinguistic approaches*, 371–388. Oxford: Oxford University Press.
Shank, Christopher, Koen Plevoets & Julie Van Bogaert. 2016. A multifactorial analysis of *that*/zero alternation: The diachronic development of the zero complementizer with *think*, *guess* and *understand*. In Jiyoung Yoon & Stefan Th. Gries (eds.), *Corpus-based approaches to Construction Grammar*, 201–240. Amsterdam & Philadelphia: John Benjamins.
Smith, Adam. 2009. Light verbs in Australian, New Zealand and British English. In Pam Peters, Peter Collins & Adam Smith (eds.), *Comparative studies in Australian and New Zealand English: Grammar and beyond*, 139–154. Amsterdam & Philadelphia: John Benjamins.
Stubbe, Maria & Janet Holmes. 1995. *You know*, *eh* and other 'exasperating expressions': An analysis of social and stylistic variation in the use of pragmatic devices in a sample of New Zealand English. *Language & Communication* 15 (1). 63–88.

Stubbs, Michael. 2002. *Words and phrases. Corpus studies of lexical semantics*. Oxford: Blackwell.
Tizón-Couto, Beatriz. 2014. *Clausal complements in native and learner spoken English. A corpus-based study with LINDSEI and VICOLSE*. Bern: Peter Lang.
Tottie, Gunnel. 1991. Conversational style in British and American English: The case of backchannels. In Karin Aijmer & Bengt Altenberg (eds.), *English corpus linguistics: Studies in honour of Jan Svartvik*, 254–271. London & New York: Longman.
Tottie, Gunnel. 2015. *Uh* and *um* in British and American English: Are they words? Evidence from co-occurrence with pauses. In Rena Torres Cacoullos, Nathalie Dion & André Lapierre (eds.), *Linguistic variation: Confronting fact and theory*, 38–54. New York & London: Routledge.
Verga, Laura & Sonja A. Kotz. 2013. How relevant is social interaction in second language learning? *Frontiers in Human Neuroscience* 7. 550. doi:10.3389/fnhum.2013.00550.
Waara, Renee. 2004. Construal, convention, and constructions in L2 speech. In Michel Achard & Susanne Niemeier (eds.), *Cognitive linguistics, second language acquisition, and foreign language teaching*, 51–75. Berlin & New York: Mouton de Gruyter.
Wible, David. 2008. Multiword expressions and the digital turn. In Fanny Meunier & Sylviane Granger (eds.), *Phraseology in foreign language learning and teaching*, 163–181. Amsterdam & Philadelphia: John Benjamins.
Wood, David. 2010. *Formulaic language and second language speech fluency: Background, evidence and classroom applications*. London: Continuum.