Unraveling Linguistic Productivity: Insights into usage, processing and variability

Ghent, May 21–23, 2024

Background information

When speakers produce or interpret language structures, they rely on a structured inventory of grammatical rules or constructions. Some of these are highly productive, having a broad domain of application and are readily available to coin new applications.

This phenomenon has long been observed in morphology. For instance, speakers of Dutch can readily apply the morphological rule Verb+*baar* to create new adjectives meaning 'that can be Verb-ed', such as in *een twitter·baar stuk tekst* 'a twitterable text chunk'. By contrast, other rules such as Verb+*(e)lijk*, as in *ondraag·lijk* 'unbearable' are not productive (Booij 2002). As a consequence, **twitter·lijk* is completely out.

But also syntactic rules and constructions can be productive to varying degrees (Zeldes 2012), since they offer slots that can be filled with a range of words, including neologisms, that is, words that have never been encountered yet in this construction. For instance, in Spanish, some inchoative auxiliary constructions apply to a seemingly infinite set of infinitives (e.g. [*empezar a* INF], *Carlos empieza a bailar '*Carlos starts to danse'), while others such as [*romper a* INF] only apply to a restricted set of infinitives. In addition, the auxiliary slot itself is considerably more productive in Spanish than in French, for instance, allowing up to 25 different verbs (Van Hulle & Enghels 2021). Similar observations hold for minimizing noun constructions that reinforce the negation of a predicate in the [*geen* N predicate] construction (e.g. *Hij komt geen meter vooruit* 'lit. he doesn't progress a meter'). In Netherlandic Dutch, no less than 244 different nouns have been attested, some of which combine with a very open list of predicates (e.g. [*geen mens*] 'no human being'), while others appear to be extremely restricted, constituting almost idioms (Van den Heede & Lauwers, f.c.).

Productivity is a theoretical construct, an abstract property of linguistic structure that forms part of the implicit knowledge speakers have about a language. Not only does it play a fundamental role in synchronic language description, it is also a crucial concept in language change (i.a. Hilpert 2013, Traugott & Trousdale 2013, Perek 2016) and language acquisition (Tomasello 2003; Yang 2016; Hartsuiker & Bernolet 2017).

However, despite all the research that has been done in these areas, productivity still raises many questions.

1° **Corpus approaches.** Traditionally, corpus linguists have inferred productivity from the sum of utterances produced by the speakers of a language, i.e. from language usage, both synchronically and diachronically. A whole bunch of measurements (Baayen 1992, 2001, 2009) such as type/token and hapax/token ratio, have been proposed, and yet new ones may still be conceivable. However, it remains to be determined **how they correlate with each other and which dimensions of productivity** they hint at (Barðdal 2008, Zeldes 2012).

2° **Psycholinguistic approaches.** Productivity is more than a set of descriptive quantities inferred from attested usage. As a dimension of what people (implicitly) know about their language(s), productivity forms a part of their mental language representations and, as such, it manifests in elicited language behavior in psycholinguistic experiments. A first question that arises in this context is **how productivity metrics based on corpora match speaker's intuitions** when it comes to assessing the availability or applicability of constructions or rules to specific lexical items. For instance, since hapaxes, the hallmark

of extensibility, are by definition low-frequent items, they are good candidates to comply with the "floor mismatch" observed by Bader & Häussler (2010) according to which rare and even unattested items may still have unexpectedly high acceptability ratings (cf. also Divjak 2017), especially with highly productive constructions. In this respect, cases of diverging acceptability of hapaxhood may be key to unravel the underlying mechanisms of productivity. Lower acceptability may also come with a processing cost, which leads to a second series of questions: for instance, **how are coinages of more or less productive constructions or rules processed in on-line adult speech**, both in **production** and **comprehension**, as measured with eye-tracking and EEG? Could these methods find evidence for differences in processing of fully schematic productivity and analogy-based local extensions (Barðdal 2008, Yang 2016), and show the impact of semantic links (for instance Barðdal 2008; Perek 2016) or semantic variability (Goldberg 2019: 65)? Finally, extensibility, or even over-extension, is a crucial factor in the process of **language learning** (Tomasello 2003; Yang 2016).

3° **Sociolinguistic approaches**. Both corpus linguists and psycholinguists tend to abstract away from the individual. However, since productivity may be seen as a constrained form of creativity (Goldberg 2019: 1; Hoffmann 2020), one may ask **to which extent productivity is determined by socio-biographical factors** (Dabrowska 2018). Thus, a third series of questions relates to individual variation and personality traits.

4° **Diachronic approaches.** Since productivity entails extensibility, diachronic linguistics may be considered a useful heuristic to better understand the process of extensibility on the basis of past extensions of the construction (e.g. Perek 2016). Therefore, this workshop particularly welcomes contributions on **changing productivity in the past** (Hartmann 2018) **and the mechanisms that triggered it**. Also, it is still unclear **how type frequency relates to changing (token) frequency,** which supposes, of course, that one is able to tease both quantities apart, which is still a matter of debate (Barðdal et al. f.c.; Feltgen 2020).

5° **Integration in the theory of grammar.** Equipped with a better understanding of what productivity is, it still remains to be seen how it can be integrated in a theory of grammar, be it in a formalist or cognitive-functionalist tradition. As to the latter, even in Construction Grammar, the natural habitat for studying productivity, it is not clear how productivity relates to schematicity, entrenchment and collocational constraints (in the case of multiple slots). In formalist approaches, partial productivity, conceived as a (weak) constraint, fuels the grammar vs. usage debate (Newmeyer 2003) and may call for the integration of probabilistic information (Francis 2022: 38-42), e.g. in Stochastic Optimality Theory. Finally, it remains to be examined how these linguistic models relate to psycholinguistic modelling. Papers that address these theoretical issues are also particularly welcome, including historiographical contributions about the history of the concept and its epistemological underpinnings. In sum, productivity, as a theoretical construct, can only be accessed when it is or has been "at work". Its many guises raise a plethora of conceptual/ontological, empirical and theoretical questions, with many methodological questions at the forefront.

This conference organized by the *The Language Productivity @ Work Consortium* particularly welcomes papers on these topics, but other related topics can also be proposed, as long as they are relevant for a better understanding of productivity.

References

Baayen, R. H. 1992. Quantitative aspects of morphological productivity. In *Yearbook of Morphology 1991*, ed. Booij, G. E., & J. Marle. Dordrecht: Kluwer Academic Publishers. 109–149.

Baayen, R. H. 2001. *Word frequency distributions*. (Text, Speech and Language Technologies 18). Dordrecht, Boston & London: Kluwer.

Baayen, R. H. 2009. Corpus linguistics in morphology: morphological productivity. In *Corpus Linguistics. An international handbook*, ed. Lüdeling, A.& M. Kyto. Berlin: De Gruyter. 900–919.

Bader M. & Häussler J. “Toward a model of grammaticality judgments “. *Journal of Linguistics*, 46/2: 273-330.

Barðdal, J. 2008. *Productivity: Evidence from Case and Argument Structure in Icelandic*. (Constructional Approaches to Language 8). Amsterdam: Benjamins.

Barðdal, J., Enghels, R., Feltgen, Q., Van Hulle, S. & Lauwers, P. f.c. “Productivity in Diachrony”. Ledgeway A., Breitbarth A., Kiss, K., Salmons J. & Simonenko A., *Wiley Blackwell Companion to Diachronic Linguistics*. Wiley-Blackwell.

Booij, G. 2002. *The Morphology of Dutch*. Oxford: Oxford UP.

Bybee, J. 2001. *Phonology and language use*. Cambridge: Cambridge UP.

Dabrowska, E. 2018. Experience, aptitude and individual differences in native language ultimate attainment. *Cognition*, *178*, 222–235.

Divjak, D. 2017. The role of lexical frequency in the acceptability of syntactic variants: evidence from that-clauses in Polish. *Cognitive Science* 41/2. 354-382.

Feltgen, Q. 2020. Diachronic Emergence of Zipf-like Patterns in Construction-Specific Frequency Distributions: A Quantitative Study of the *Way Too* Construction. *Lexis* [Online], 16 [URL : http://journals.openedition.org/lexis/4968].

Feltgen, Q. 2020. “Diachronic Emergence of Zipf-like Patterns in Construction-Specific Frequency Distributions: A Quantitative Study of the *Way Too* Construction”, *Lexis* [Online], 16.

Firth, J.R. 1957. *A synopsis of linguistic theory 1930-1955*. *Studies in Linguistic Analysis*. Oxford : Philological Society, 1-32.

Francis, E. 2022. *Gradient acceptability and linguistic theory*. Oxford: Oxford University Press.

Goldberg, A. 2019. *Explain me this. Creativity, Competition, and the Partial Productivity of Constructions.* Princeton: Princeton University Press.

Hartmann, Stefan. 2018. “Derivational Morphology in Flux: A Case Study of Word-Formation Change in German.” Cognitive Linguistics, 29(1): 77–119.

Hartsuiker, R. J., & S. Bernolet. 2017. The development of shared syntax in second language learning. *Bilingualism: Language and Cognition* 20(2): 219–234.

Hilpert, M. 2013. *Constructional Change in English .Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: Cambridge UP.

Hoffmann, T. 2018. "Creativity and Construction Grammar: Cognitive and Psychological Issues". *Zeitschrift für Anglistik und Amerikanistik*, 66/3: 259-276.

Newmeyer, F. 2003. "Grammar is Grammar and Usage is Usage." *Language*, 79/4, 682-707.

Perek, F. 2016. Using distributional semantics to study syntactic productivity in diachrony: A case study. *Linguistics* 54(1): 149–188.

Suttle, L. & A. Goldberg. 2011. The partial productivity of constructions as induction, *Linguistics* 49/6: 1237-1269.

Tomasello, M. 2003. *Constructing A Language: A Usage Based Theory of Language Acquisition*. Harvard: UP.

Traugott, E. & G. Trousdale. 2013. *Constructionalization and constructional change*. Oxford: Oxford UP.

Van Hulle, S. & Enghels, R. 2021. “De Spaanse inchoatiefconstructie in beeld: clusteranalyse als antwoord op het quasi-synonymie vraagstuk. *Handelingen (KZM),* 75/1: 277-305.

Van den Heede, M. & Lauwers, P. f.c. “Syntactic productivity under the microscope: the lexical and semantic openness of Dutch minimizing constructions". *Folia Linguistica*.

Van Wettere, N. 2021. “Productivity of French and Dutch (semi-)copular constructions and the adverse impact of high token frequency”. International Journal of Corpus linguistics 26: 396-428.

Yang, C. 2016. *The price of productivity*. MIT press.

Zeldes, A. 2012. Productivity *in Argument Selection: From Morphology to Syntax.* Berlin: De Gruyter.