

General introduction:

A comparative perspective on probabilistic variation in grammar

Abstract

This special collection brings together research exploring and evaluating probabilistic variation patterns from a comparative perspective, thus highlighting current work situated at the crossroads of research on usage-based theoretical linguistics, variationist linguistics, and sociolinguistics. The contributions in the collection advance our understanding of the plasticity of syntactic knowledge on the part of language users with diverse regional and/or cultural backgrounds, and demonstrate how a probabilistic approach to grammatical variation can offer insight into the scope and limits of language variation. In this general introduction to the special collection, we provide some essential background for the sake of perspectivization, and subsequently summarize the contributions in the collection.

1. Background

Work that comes under the remit of this special collection (a) acknowledges that variation between different ways of saying the same thing is sensitive to multiple and sometimes competing constraints which influence linguistic choice-making in subtle, probabilistic ways; (b) is specifically concerned with grammatical variation; and (c) explores how probabilistic choice making processes differ across varieties of the same language.

Why is scholarship along these lines important? A sizable body of research indicates that probabilistic patterns and mechanisms are pervasive on all levels of language (see, e.g., the contributions in Bod, Hay & Jannedy 2003a). As to grammar specifically, we know that intra-systemic grammatical variation – that is, variation within and across varieties of the same language – is highly systematic, and that the determinants of this variation are numerous, multifactorial, and probabilistically conditioned (e.g. Gries 2003; Bresnan & Hay 2008; Tagliamonte, Durham & Smith 2014; Szmrecsanyi et al. 2016). Results of such studies are generally taken to be evidence for a model of grammar that is quantitative and probabilistic. Those probabilistic approaches to language variation that are explicitly usage-based in nature are additionally committed to the notion that grammar is the “cognitive organization of one’s experience with language” (Bybee 2006:711). Hence, variation patterns are thought to be learned directly, perhaps entirely, from an individual’s exposure to other speakers’ language (Bybee & Hopper 2001; Bresnan & Ford 2010). From such a perspective, the grammar is inherently variable, as successive generations of speakers are exposed to sets of exemplars that differ in subtle ways. This variation may be shaped by social, cognitive or functional factors, whose influences on individual speakers’ production (and comprehension) are aggregated to result in population-level linguistic phenomena. To the extent that the linguistic experience of different speakers or communities varies, we expect gradient differences in the grammar to emerge as speakers adjust their knowledge of linguistic phenomena to match that of their input. Needless to say, this perspective differs sharply from rule-based approaches which assume that grammatical knowledge is categorical, possibly biologically innate, and that linguistic variation is theoretically irrelevant to the investigation of the principles that determine syntactic structure. Note also that

while rule-based approaches are interested in the categorical (un-)grammaticality of a given linguistic form, probabilistic approaches challenge this categoriality and instead assume a “cline of well-formedness” (Bod, Hay & Jannedy 2003b:4)

Against this backdrop, current theorizing about the nature of grammatical knowledge is often framed as an opposition between two diametrical poles: fully usage-based (i.e. exemplar-based) (e.g. Pierrehumbert 2006) versus rule-based approaches (e.g. Chomsky & Halle 1968). While the respective approaches involve very different assumptions, each has had considerable empirical success over the years. Still, each approach is not without its limitations, and there is an emerging view that a kind of hybrid model is necessary in order to account for the full range of language phenomena (see Guy 2014 for discussion). In any event, both usage- and rule-based models of grammar are mentalistic, in that they view language as a cognitive object, and it is this common ground upon which most hybrid models are founded.

One hybrid approach that we would like to discuss here in a bit more detail by way of exemplification is the variation-centered, usage- and experience-based probabilistic grammar approach developed by Joan Bresnan and collaborators (e.g. Bresnan 2007; Bresnan & Hay 2008; Bresnan & Ford 2010). Work in this particular school of thought makes two key assumptions (in addition to those mentioned at the beginning of this section): First, grammatical *knowledge* has a probabilistic component, and language users have powerful predictive capabilities. Second, this probabilistic knowledge is derived in large part from language experience, and so is subtly, but dynamically (re)constructed throughout speakers’ lives. The probabilistic nature of grammar is supported by evidence showing that the likelihood of finding a particular linguistic variant in a particular context in a corpus corresponds to the intuitions that speakers have about the acceptability of the variants (see Bresnan & Ford 2010; Klavan & Divjak 2016). Bresnan (2007:76–84), for example, used a scalar rating task based on corpus materials (transcriptions of spoken dialogue passages) as stimuli to model subjects’ responses regarding the naturalness of dative variants in contexts. These responses were compared to the predictions of the dative alternation regression model reported in Bresnan et al. (2007). It turned out that subjects’ gradient (i.e. probabilistic) naturalness ratings overlapped significantly with corpus-generated probabilities; hence, speakers’ implicit knowledge about language must be to some extent probabilistic in nature.

Bresnan-style probabilistic grammar work is a hybrid approach because it emphasizes the association of conventional rules or constraints with probabilities learned from experience – in other words, we are dealing with a “balanced diet” (Guy 2014:65) model of syntax enriched with both qualitative and quantitative aspects. We would like to stress in connection that the methodologies and research questions in this research tradition are largely compatible with work in modern variationist sociolinguistics (see, e.g., Labov 1982). Specifically, what takes center stage in both schools is how and why (i.e. subject to which constraints, be they language-internal or language-external) people choose between “alternate ways of saying ‘the same’ thing” (Labov 1972: 188). Indeed, variationist sociolinguists have been in the business of analyzing variation patterns probabilistically for decades (consider e.g. Cedergren & Sankoff 1974), and one way of conducting probabilistic sociolinguistics that is particularly pertinent to the present Special Collection is to compare community grammars via the Comparative Sociolinguistics method (see Tagliamonte 2001), which is designed to assess the extent to which variation patterns across similarities provide a signal of (historical) relatedness.

Whatever the actual sub-disciplinary flavor, most probabilistic approaches to analyzing variation tend to be inherently usage-based in that they incorporate statistical regularities derived from experience, yet they associate these quantitative patterns not (only) with surface forms or lexical items (as in pure exemplar models), but with abstract features or constraints (e.g. whether a constituent refers to an animate entity or not). Many, though by no means all of these features are taken to represent inherent, universal biases in language structure, e.g. the overwhelming tendency to map animate referents to more prominent positions (e.g. Rosenbach 2005), or the tendency in VO languages such as English to place lighter (shorter) elements before heavier (longer) ones (e.g. Wasow & Arnold 2003). Probabilistic accounts can account for gradient, experience-driven variability within the context of universal constraints on the range of possible variation. Probabilistic accounts thus share much in common with Optimality Theory (Kager 1999), the difference being that probabilistic grammars do not assume a fixed set of innate constraints. From the perspective of probabilistic linguistics, variation (inter- and intra-systemic) arises from the interplay between biases in language production and comprehension and acquired syntax-semantic associations, and this interplay leads to statistical variability in the distribution of forms which speakers implicitly learn (see MacDonald 2013). On the empirical plane, probabilistic variation analysis tends to be based more often than not on the analysis of naturalistic corpus data (this includes collections of sociolinguistic interviews). As we have seen, this is sometimes supplemented by experimental approaches (see e.g. Rosenbach, this volume).

Importantly, social meaning and socially conditioned variation (including regional differentiation) is entirely compatible with – even predicted by – probabilistic grammar models. Community-specific social forces, e.g. language attitudes or stylistic preferences, undoubtedly shape biases in individual speakers' production and comprehension, while at the same time, ad hoc meaning formation that arises during individuals' interactions can lead to innovation and greater variability among syntactic forms and their semantic cues. The resulting patterns are in turn reflected in specific forms' distributions across different social groups/contexts. Probabilistic grammar models are thus also consonant with the view of the developing subfield of Cognitive Sociolinguistics (see e.g. Geeraerts, Kristiansen & Peirsman 2010), which seeks to more fully integrate both the social and cognitive dimensions into more complete models of language structure and variation.

If, as usage-based approaches generally argue (Bybee & Hopper 2001; Scott-Philips & Kirby 2010), individual-level behavior leads to population-level language patterns, and if individual behaviors are guided in part by universal cognitive processes, several predictions follow: (a) The influence of certain cognitive factors on quantitative syntactic variation in across different (sub)varieties of a given language should be relatively stable in terms of the direction of those factors' influence. (b) Subtle variation in the types and frequencies of constructions will lead to gradient, yet detectable differences in the strength of different factors' influence on speakers' syntactic choices. (c) This variation in the use of specific constructions may be driven by stylistic preferences among registers or speakers, by situational forces such as language/dialect contact, by cognitive pressures related to language processing, or by normal dialectal drift.

As we saw in the foregoing discussion, comparative probabilistic grammar analysis as defined at the beginning of this section (attention to multiple probabilistic constraints; focus on syntactic/grammatical variation; interest in contrasts between language varieties) has sure enough been around for a while, not only in the well-known guise of comparative (variationist)

sociolinguistics (see, e.g., Jones & Tagliamonte 2004; Tagliamonte & Smith 2005; Tagliamonte 2014, among many other studies), but also in other schools of variation study (e.g. Bresnan & Hay 2008; Bresnan & Ford 2010; Ehret, Wolk & Szmrecsanyi 2014; Gries & Deshors 2014; Hinrichs, Szmrecsanyi & Bohmann 2015; Wulff, Lester & Martinez-Garcia 2014). But it is only in recent years that the predictions outlined in the previous paragraph have begun to be explored more systematically. We take the liberty to illustrate this trend by sketching a research project (2013-2021) based at the KU Leuven and entitled “Exploring probabilistic grammar(s) in varieties of English around the world”, which investigates three syntactic alternations (see (1) – (3)) in some nine international varieties of English: British English, Canadian English, Irish English, New Zealand English, Hong Kong English, Indian English, Jamaican English, Philippine English, and Singapore English (Szmrecsanyi, Grafmiller, Heller & Röthlisberger 2016).

- (1) The dative alternation (see Röthlisberger, Grafmiller & Szmrecsanyi 2017)
 - a. *I'd given Heidi my T-Shirt* (the ditransitive dative variant)
 - b. *And I'd given the key to Helen* (the prepositional dative variant)
- (2) The genitive alternation (see Heller, Szmrecsanyi & Grafmiller 2017)
 - a. *the country's economic crisis* (the *s*-genitive)
 - b. *the economic growth of the country* (the *of*-genitive)
- (3) The particle placement alternation (see Grafmiller & Szmrecsanyi to appear)
 - a. *you can just cut the tops off* (verb-object-particle order)
 - b. *cut off the flowers* (verb-particle-object order)

Some key findings include the following. First, probabilistic grammars are on the whole surprisingly stable in a cross-variety perspective. Specifically, we very rarely see reversals in effect directions: constraints tend to have the same qualitative effect across varieties, which points to a fairly stable core probabilistic grammar. In more quantitative terms, Heller (in preparation) calculates various core grammar-hood coefficients, which can range between 0 (no probabilistic similarity whatsoever between the varieties under study) and 1 (maximal probabilistic similarity) and finds that in the case of the English genitive alternation, coefficients range between approximately 0.6 and 0.9. However, there do seem to be interesting quantitative differences with regard to the *effect size* of the constraints on variation. These quantitative differences we tend to find only in those contexts where neither alternate is more or less difficult to process (Szmrecsanyi et al. 2016:132), and where shifting usage frequencies in language-internal variation may have led to regional differences between users' probabilistic grammars (Röthlisberger, Grafmiller & Szmrecsanyi 2017). Curiously, one of the constraints that is malleable fairly consistently across varieties and alternations turns out to be constituent length. Constituent length fuels the principle of end-weight (Behaghel 1909; Wasow & Arnold 2003) – in VO languages such as English, language users tend to place longer, heavier constituents after shorter, lighter ones – and is often thought to be rooted in the architecture of speech processing system (e.g. Hawkins 1994). Precisely because of this rootedness, end-weight is not a prime suspect for probabilistic cross-variety contrasts, but the data suggest otherwise. Second, from a dialect-typological point of view varieties often pattern along native versus non-native (or Inner Circle versus Outer Circle; see Kachru 1992) lines. For example, Heller et al. (2017) show that the well-known animacy constraint on genitive variation – animate possessors attract the *s*-genitive, rather than the *of*-genitive – is stronger in native varieties of English (e.g. British English) than in non-native varieties of English (e.g. Indian English). Similarly, Röthlisberger et al. (2017) and Grafmiller & Szmrecsanyi (to appear) find that some of the largest deviations in individual factor effects on the

dative alternation and particle placement alternations respectively occur in the non-native varieties. Third, in an alternation-oriented (or: variable-oriented) perspective different alternations differ as to how amenable they are to probabilistic indigenization. “Probabilistic indigenization” Szmrecsanyi et al. (2016:133) define as the process “whereby stochastic patterns of internal linguistic variation are reshaped by shifting usage frequencies in speakers of post-colonial varieties”.¹ What seems to be the case is that the less abstract a given syntactic alternation is and the more lexical slots it has, the more likely it is to exhibit cross-varietal probabilistic indigenization effects. This is why the particle placement alternation is quite variable in a cross-variety perspective, while the genitive alternation (which is an almost purely positional alternation) is not (see Szmrecsanyi et al. 2016:133 for more discussion).

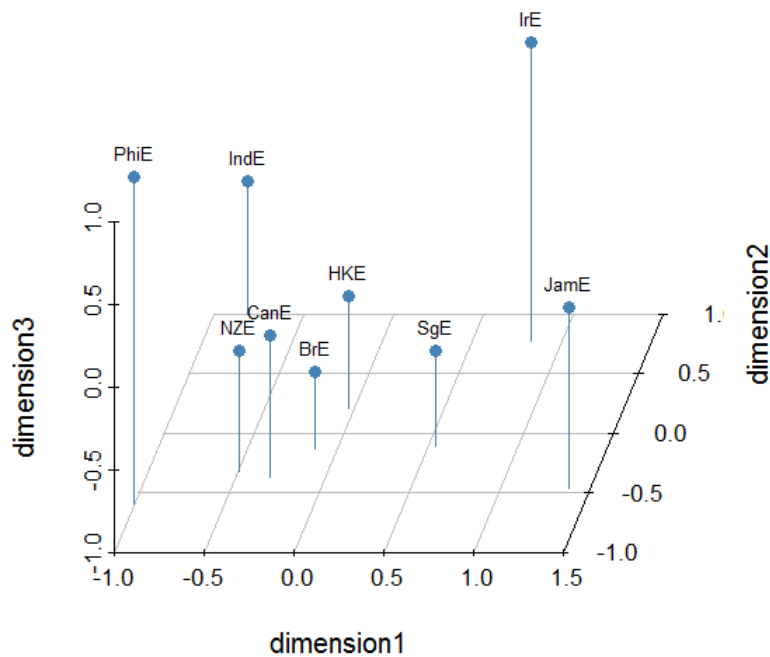


Figure 1: Probabilistic distances between nine varieties of English, based on regression analysis of the English dative, genitive, and particle placement alternations. Distance between data points in the plot is proportional to aggregate probabilistic distances.

¹ Röthlisberger et al. (2017) tighten the concept of probabilistic indigenization to "cognitive indigenization", to stress the outcome of probabilistic indigenization, i.e. the emergence of distinct lects.

In the bird's eye view, aggregate probabilistic grammar distances between the varieties may be visualized as follows. We begin by using coefficient estimates of by-variety, by-alternation regression models to create a series of Euclidean distance matrices (one for each alternation under analysis), in which pairwise distances between varieties calculate as the square root of all coefficient differentials (see Röthlisberger 2018:78-83 for details). In a second step, alternation-specific distance matrices are combined to generate a synoptic distance matrix. Third, we use Multidimensional Scaling (MDS) to reduce the synoptic distance matrix to a lower-dimensional representation, as in Figure 1. The pattern that emerges can be summarized as follows. Most native varieties (British English, Canadian English, and New Zealand English) cluster at the bottom of the cube. Irish English is quite different, in a probabilistic grammar perspective, from the other native varieties, maybe thanks to its Celtic substrate. On the other hand, Singapore English and Hong Kong English are two non-native varieties which are remarkably close to many of the native varieties in the sample (we note, in this connection, that Singapore English is often argued to be in the process of becoming a genuine native variety (Leimgruber 2013:122). Indian English, Jamaican English, and Philippines English are, each in its own way, dissimilar probabilistically from the other varieties in the sample.

It is precisely against the backdrop of research in this probabilistic-cum-comparative spirit that the present special collection has been designed.

2. Contributions in the special collection

The contributions in this special collection represent new steps toward a better understanding of the nature and limits of grammatical variation. At the same time, they also broaden the scope of probabilistic grammar research in a number of important ways. The collection thus highlights the healthy diversity of perspectives on probabilistic variation patterns from a comparative perspective that we find in variation studies today.

In her contribution entitled “Constraints in contact: Animacy in English and Afrikaans genitive variation – a cross-linguistic perspective”, Anette Rosenbach undertakes a comparative analysis of the effect of possessor animacy on genitive variation in British English, Afrikaans, and South African English, three languages/varieties with structurally very similar genitive variation grammars. Study 1 shows, on the basis of an analysis of a parallel corpus data, that the Afrikaans prenominal possessive (the *se*-genitive, as in *Harry se hart*) is less strongly attracted to animate possessors than its cousin, the English *s*-genitive (as in *Harry's heart*). Study 2 marshals a forced-choice experiment to demonstrate that the weaker animacy constraint in Afrikaans carries over to the L2 English of L1 Afrikaans speakers. Rosenbach concludes that English (and varieties of English) partake in a typological continuum of possession splits according to possessor animacy, and that probabilistic constraint strengths may be transferred in contact situations. What is stable is a pattern known as harmonic alignment in the literature: language users tend to place animate possessors first.

Jeroen Claes ("Probabilistic grammar: The view from Cognitive Sociolinguistics") advocates drawing inspiration from cognitive (socio)linguistics for the sake of defining theoretically better motivated predictor/constraint sets for probabilistic grammar analysis. Cognitive Linguistics is a theoretical orientation which posits that linguistic knowledge derives from usage, and is committed to describing language in terms of what is known from other cognitive disciplines about the functioning of the mind; cognitive *sociolinguistics* is additionally interested in how social and cultural factors shape

linguistic awareness, cognition, and usage. To highlight the theoretical benefits of marrying cognitive (socio)linguistics to probabilistic grammar research, Claes presents three case studies (variable agreement with existential *haber* in three varieties of Spanish, variable agreement with existential *there be* in British English, and subject pronoun expression in Cuban Spanish), and demonstrates how a set of cognitively motivated constraints (markedness of coding, statistical preemption, and structural priming) can help us to better understand grammatical variation patterns.

Claire Childs, in her contribution entitled "Integrating syntactic theory and variationist analysis: The structure of negative indefinites in regional dialects of British English", presents a comparative sociolinguistics analysis of variation between *not*-negation, *no*-negation, and negative concord (a.k.a. multiple negation) in three UK communities (Glasgow, Tyneside, and Salford). The aim of the analysis is to assess the explanatory power of two competing formal accounts for the variation under study. These accounts make different predictions about the distributions of variants, and it these distributions that are checked in the variationist analysis. Investigating conditioning factors such as verb type, verb phrase complexity, and discourse status, Childs shows that *no*-forms are marked syntactically for negation Determiner Phrase-internally in cases of *no*-negation (as well as in pre-verbal position and fragment answers), but not in negative concord. This evidence supports an account positing that *no*-negation is derived via negative-marking within the Determiner Phrase followed by movement to the Negative Phrase for sentential scope.

The contribution by Natalia Levshina ("Communicative efficiency and probabilistic grammar: Bayesian mixed-effect regression models of *help* + (*to*) Infinitive in varieties of web-based English") builds bridges to information theory and is particularly interested in the role that information content and predictability play in the choice between competing complementation variants after the English verb *help* (other factors subject to study include cognitive complexity, horror aequi, and iconicity). Variation is studied in a number of World Englishes (Australian English, Ghanaian English, British English, Hong Kong English, Indian English, Jamaican English, and US American English). Levshina marshals Bayesian regression analysis to model the interplay of the constraints on variation, and demonstrates that the more explicit complementation variant with *to* (as in *Mary helped John to cook the dinner*) is particularly favored in contexts with high information content. What is more, she observes a stable, universal pattern of communicatively efficient behavior in the probabilistic grammars of all the varieties under study.

In "Spoken syntax in a comparative perspective: the dative and genitive alternation in varieties of English", Benedikt Szmrecsanyi, Jason Grafmiller, Joan Bresnan, Anette Rosenbach, Sali Tagliamonte, and Simon Todd introduce newly created and freely available datasets. These were compiled from previously analyzed data and are designed to facilitate the quantitative investigation of syntactic variation in spoken language from a comparative perspective. The datasets cover the genitive (*anthropology's history* versus *the history of anthropology*) and dative (*give me some pizza* versus *give some pizza to me*) alternations in four vernacular varieties of English: American English, British English, Canadian English, and New Zealand English. To highlight the potential of the data source, the authors conduct a pilot study that suggests on the one hand that while there are a number of subtle probabilistic contrasts between the regional varieties under study, we see overall a striking degree of cross-varietal stability and homogeneity. On the other hand, the authors find it surprisingly hard to replicate probabilistic contrasts reported in the previous literature, which raises questions about the generalizability of results in contrastive probabilistic grammar research.

Taken together, then, the contributions in this special collection push forward the theoretical and empirical state of the art in comparative probabilistic grammar analysis in a number of ways. As to the variation phenomena subject to study, the contributions look beyond the usual suspects (i.e. the English genitive and dative alternations) and cover lesser studied grammatical variation phenomena such as agreement patterns and subject expression (Claes), negative indefinites (Childs), and verbal complementation (Levshina). These variation phenomena are investigated in a range of languages (English, Spanish, Afrikaans) and language varieties some of which one does not often see in the variationist/probabilistic grammar literature (consider e.g. Afrikaans, South African English, or Ghanaian English). On the technical side, the contributions deploy a variety of analysis techniques: beside run-of-the-mill regression analysis, we find parallel corpus study and forced-choice experiments (Rosenbach), as well as Bayesian mixed-effect regression analysis (Levshina). Szmrecsanyi et al. essentially conduct a meta-study to evaluate, among other things, the replicability of probabilistic grammatical contrasts. Last but not least, the contributions cross disciplinary boundaries in various ways: Rosenbach explores the interface between probabilistic grammar research and cross-linguistic typology; Claes highlights crosspollination potential between variationist/probabilistic linguistics and cognitive (socio)linguistics; Childs shows how formal syntactic theory may inspire variationist linguistics and vice versa; and Levshina taps into information theory to derive predictors for variationist analysis.

References

- Behaghel, Otto. 1909. Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. *Indogermanische Forschungen* 25. 110–142.
- Bod, Rens, Jennifer Hay & Stefanie Jannedy (eds.). 2003a. *Probabilistic linguistics*. Cambridge, {MA}: {MIT} Press.
- Bod, Rens, Jennifer Hay & Stefanie Jannedy (eds.). 2003b. Introduction. *Probabilistic Linguistics*, 1–10. Cambridge, MA: MIT Press.
- Bresnan, Joan. 2007. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In Sam Featherston & Wolfgang Sternefeld (eds.), *Roots: Linguistics in Search of Its Evidential Base*, 75–96. Berlin: Mouton de Gruyter.
- Bresnan, Joan, Anna Cueni, Tatiana Nikitina & Harald Baayen. 2007. Predicting the Dative Alternation. In Gerlof Boume, Irene Krämer & Joost Zwarts (eds.), *Cognitive Foundations of Interpretation*, 69–94. Amsterdam: Royal Netherlands Academy of Science.
- Bresnan, Joan & Marilyn Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86(1). 168–213. doi:10.1353/lan.0.0189.
- Bresnan, Joan & Jennifer Hay. 2008. Gradient grammar: An effect of animacy on the syntax of give in New Zealand and American English. *Lingua* 118(2). 245–259. doi:10.1016/j.lingua.2007.02.007.
- Bybee, Joan. 2006. From Usage to Grammar: The Mind's Response to Repetition. *Language* 82(4). 711–733.
- Bybee, Joan & Paul Hopper. 2001. *Frequency and the Emergence of Linguistic Structure*. Amsterdam: Benjamins.
- Cedergren, Henrietta J. & David Sankoff. 1974. Variable Rules: Performance as a Statistical Reflection of Competence. *Language* 50(2). 333. doi:10.2307/412441.
- Chomsky, Noam & Morris Halle. 1968. *The sound pattern of English*. Cambridge, Mass.: MIT Press.

- Ehret, Katharina, Christoph Wolk & Benedikt Szmrecsanyi. 2014. Quirky quadratures: on rhythm and weight as constraints on genitive variation in an unconventional data set. *English Language and Linguistics* 18(02). 263–303. doi:10.1017/S1360674314000033.
- Geeraerts, Dirk, Gitte Kristiansen & Yves Peirsman (eds.). 2010. *Advances in cognitive sociolinguistics*. Berlin/New York: De Gruyter Mouton.
https://wwwling.arts.kuleuven.be/qlvl/iprints/Geeraerts_kristiansen_peirsman_2010final_Advances_Cognitive_Sociolinguistics.pdf.
- Grafmiller, Jason & Benedikt Szmrecsanyi. in preparation. Mapping out particle placement in Englishes around the world: A case study in probabilistic indigenization.
- Gries, Stefan Th. 2003. *Multifactorial Analysis in Corpus Linguistics: A Study of Particle Placement*. New York: Continuum Press.
- Gries, Stefan Th. & Sandra C. Deshors. 2014. Using regressions to explore deviations between corpus data and a standard/target: two suggestions. *Corpora* 9(1). 109–136.
doi:10.3366/cor.2014.0053.
- Guy, Gregory R. 2014. Linking usage and grammar: Generative phonology, exemplar theory, and variable rules. *Lingua* 142. (SI: Usage-Based and Rule-Based Approaches to Phonological Variation). 57–65. doi:10.1016/j.lingua.2012.07.007.
- Hawkins, John A. 1994. *A performance theory of order and constituency*. Cambridge; New York: Cambridge University Press.
- Heller, Benedikt. in preparation. The genitive alternation in varieties of English. Leuven: KU Leuven PhD dissertation.
- Heller, Benedikt, Benedikt Szmrecsanyi & Jason Grafmiller. 2017. Stability and Fluidity in Syntactic Variation World-Wide: The Genitive Alternation Across Varieties of English. *Journal of English Linguistics* 45(1). 3–27. doi:10.1177/0075424216685405.
- Hinrichs, Lars, Benedikt Szmrecsanyi & Axel Bohmann. 2015. Which-hunting and the Standard English relative clause. *Language* 91(4). 806–836. doi:10.1353/lan.2015.0062.
- Jones, Megan & Sali Tagliamonte. 2004. From Somerset to Samaná: Preverbal did in the voyage of English. *Language Variation and Change* 16. 93–126.
- Kachru, Braj B. (ed.). 1992. *The Other tongue: English across cultures*. 2nd ed. (English in the Global Context). Urbana: University of Illinois Press.
- Kager, Rene. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
doi:10.1017/CBO9780511812408. <http://ebooks.cambridge.org/ref/id/CBO9780511812408>
(1 June, 2017).
- Klavan, J. & D. Divjak. 2016. The cognitive plausibility of statistical classification models: Comparing textual and behavioral evidence. *Folia Linguistica* 50(2). 355–384.
doi:<https://doi.org/10.1515/flin-2016-0014>.
- Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45. 715–762.
- Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Philadelphia Press.
- Labov, William. 1982. Building on empirical foundations. In Winfred Lehmann & Yakov Malkiel (eds.), *Perspectives on Historical Linguistics*, 17–92. Amsterdam, Philadelphia: Benjamins.
- Leimgruber, Jakob R. E. 2013. *Singapore English: structure, variation, and usage*. Cambridge: Cambridge University Press.
- MacDonald, Maryellen C. 2013. How language production shapes language form and comprehension. *Frontiers in Psychology* 4. 1–16. doi:10.3389/fpsyg.2013.00226.
- Pierrehumbert, Janet B. 2006. The next toolkit. *Journal of Phonetics* 34(4). 516–530.
doi:10.1016/j.wocn.2006.06.003.

- Rosenbach, Anette. 2005. Animacy versus weight as determinants of grammatical variation in English. *Language* 81(3). 613–644.
- Röthlisberger, Melanie. 2018. *Regional variation in probabilistic grammars*. Leuven: KU Leuven. PhD dissertation.
- Röthlisberger, Melanie, Jason Grafmiller & Benedikt Szmrecsanyi. 2017. Cognitive indigenization effects in the English dative alternation. *Cognitive Linguistics* 28(4). doi: 10.1515/cog-2016-0051.
- Scott-Phillips, Thomas C. & Simon Kirby. 2010. Language evolution in the laboratory. *Trends in Cognitive Sciences* 14(9). 411–417. doi:10.1016/j.tics.2010.06.006.
- Szmrecsanyi, Benedikt, Jason Grafmiller, Benedikt Heller & Melanie Röthlisberger. 2016. Around the world in three alternations: Modeling syntactic variation in varieties of English. *English World-Wide* 37(2). 109–137. doi:10.1075/eww.37.2.01szm.
- Tagliamonte, Sali. 2001. Comparative sociolinguistics. In Jack Chambers, Peter Trudgill & Natalie Schilling-Estes (eds.), *Handbook of Language Variation and Change*, 729–763. Malden and Oxford: Blackwell.
- Tagliamonte, Sali A. 2014. A comparative sociolinguistic analysis of the dative alternation. In Rena Torres-Cacoullos, Nathalie Dion & André Lapierre (eds.), *Linguistic variation: Confronting fact and theory*, 297–318. London, New York: Routledge.
- Tagliamonte, Sali A., Mercedes Durham & Jennifer Smith. 2014. Grammaticalization at an early stage: future be going to in conservative British dialects. *English Language and Linguistics* 18(01). 75–108. doi:10.1017/S1360674313000282.
- Tagliamonte, Sali & Jennifer Smith. 2005. No momentary fancy! The zero in English dialects. *English Language and Linguistics* 9(02). 289–309. doi:10.1017/S1360674305001644.
- Wasow, Thomas & Jennifer Arnold. 2003. Post-verbal constituent ordering in English. In Günter Rohdenburg & Britta Mondorf (eds.), *Determinants Of Grammatical Variation in English*, 119–154. Amsterdam: Mouton de Gruyter.
- Wulff, Stefanie, Nicholas Lester & Maria T. Martinez-Garcia. 2014. That-variation in German and Spanish L2 English. *Language and Cognition* 6(02). 271–299. doi:10.1017/langcog.2014.5.