Alice Blumenthal-Dramé*

What corpus-based Cognitive Linguistics can and cannot expect from neurolinguistics

DOI 10.1515/cog-2016-0062

Received May 28, 2016; revised August 25, 2016; accepted August 25, 2016

Abstract: This paper argues that neurolinguistics has the potential to yield insights that can feed back into corpus-based Cognitive Linguistics. It starts by discussing how far the cognitive realism of probabilistic statements derived from corpus data currently goes. Against this background, it argues that the cognitive realism of usage-based models could be further enhanced through deeper engagement with neurolinguistics, but also highlights a number of common misconceptions about what neurolinguistics can and cannot do for linguistic theorizing.

Keywords: corpus-based Cognitive Linguistics, neurolinguistics, cognitive realism, levels of analysis

1 Introduction

A guiding assumption of corpus-based Cognitive Linguistics has been that linguists' statistical generalizations over corpus data can provide insights into the cognition of an idealized average language user. The view that probabilistic generalizations over corpus data to some extent reflect language users' knowledge of language is illustrated, for example, by Bresnan and Ford (2010: 205), who argue that "implicit probabilistic knowledge [...] supports, or perhaps constitutes, the language faculty." Likewise, Gries and Ellis state that

[...]language learners do not consciously tally [...] corpus-based statistics. The frequency tuning under consideration here is computed by the learner's system automatically during language usage. The statistics are implicitly learned and implicitly stored [...]; learners do not have conscious access to them. Nevertheless, every moment of language cognition is informed by these data [...]. (Gries and Ellis 2015: 241)

The last decade has seen a growing number of experimental studies confirming principled correlations between statistical generalizations over corpus data and subjects' experimental behaviors at various levels of language description,

*Corresponding author: Alice Blumenthal-Dramé, Freiburg Institute for Advanced Studies, Albert-Ludwigs-Universität Freiburg, Freiburg im Breisgau, Germany,
E-mail: alice.blumenthal@anglistik.uni-freiburg.de

including phonetics (Arnon and Cohen Priva 2013; Tily et al. 2009; Kleinschmidt and Jaeger 2015), morphology (Blumenthal-Dramé 2012; Blumenthal-Dramé et al. under revision; Schmidtke et al. 2015; Feldman et al. 2015; for review see Amenta and Crepaldi 2012), lexical processing (Frisson et al. 2005; McDonald and Shillcock 2003; Smith and Levy 2013), semantics (Griffiths et al. 2007; Jones and Mewhort 2007), multi-word sequences (Arnon and Snider 2010; Kapatsinski and Radicke 2009; Siyanova-Chanturia et al. 2011; Snider and Arnon 2012; Tremblay and Tucker 2011; Tremblay et al. 2011), syntax (Bresnan 2007; Bresnan and Ford 2010; Demberg and Keller 2008; Fine et al. 2013; Jaeger and Snider 2013; Kamide 2012; Levy 2008), and associations between words and grammatical structures (Arai and Keller 2013; Divjak 2016; Gahl and Garnsey 2004; Gries et al. 2005, 2010; Hare et al. 2007; Linzen and Jaeger 2016; Wilson and Garnsey 2009) (for a discussion of cases where experimental behaviors and corpus-derived predictions do not align, cf. Dąbrowska this issue).

The aim of this paper is not to belittle the merits of these studies – on the contrary, they have contributed to putting frameworks that conceive of language structure as gradient, probabilistic and malleable on a firmer footing, in particular probabilistic, exemplar-based and usage-based models (Bybee 2010; Goldberg 2006; Bresnan and Hay 2008; Divjak and Gries 2012; Gahl and Yu 2006; Gries and Divjak 2012; Hay and Baayen 2005). Without doubt, these studies have also yielded cognitively highly relevant insights, for example by comparing the predictive power of different metrics at different levels of language description, by conducting experiments tapping into different facets of language representation (e. g., comprehension, production, grammaticality judgements), and by demonstrating that language users' behavior is extremely sensitive to and constantly (re-)shaped under the influence of other people's linguistic behavior.

Moreover, they have brought linguistic theorizing into closer alignment with contemporary developments in the cognitive sciences, in particular probabilistic approaches to human cognition, which are increasingly used to model different areas of cognition such as causal learning, visual perception, motor control, and inferential reasoning (Chater et al. 2006; Chater and Oaksford 2008; Griffiths and Tenenbaum 2006; Griffiths et al. 2010; Tenenbaum et al. 2011; Perfors et al. 2011).

Rather, what the present contribution aims to do is to discuss how far the cognitive realism of probabilistic statements derived from corpus data currently goes (Section 2), to outline ways in which the cognitive realism of usage-based models could be further enhanced through deeper engagement with neurolinguistics (Section 3), and to set straight a number of misconceptions about what neurolinguistics can and cannot do for linguistic theorizing (Section 4). The concluding section (Section 5) summarizes the role that I expect neurolinguistics to take on in the cognitive linguistics enterprise.

2 How cognitively realistic are corpus-derived statistical models?

In the present section, I will define the sense in which I take corpus-derived statistical models to be cognitively realistic. In a first step, I will explain why I do not consider it theoretically fruitful to assume that the “exemplar store”, which is a popular way of modeling how statistical generalizations over usage data emerge in the cognitive system, actually models what goes on in people’s brains. In a second step, I will attempt to define, in positive terms, the extent to which the exemplar store can be taken to be cognitively realistic.

Usage-based models view the exemplar store as a multi-dimensional memory space over which language users compute statistical generalizations over usage data. All exemplars encountered in language use, no matter how complex, are assumed to be stored in high-resolution format, with memory traces encoding fine-grained linguistic and non-linguistic facets of the input (e. g., detailed phonetic realization, gender of the speaker, situational context, etc.). The exemplar store is continuously updated as a result of on-going experience with language, and abstractions do not supersede the individual episodic traces from which they emerge (Docherty and Foulkes 2014).

Although there has been some debate around the precise degree of granularity of exemplar traces, current models largely proceed from the working hypothesis that the memory store records perceived input in a relatively faithful fashion. This stance is summarized by Docherty and Foulkes, who claim that

[w]ith relatively few exceptions, accounts of the exemplar model by both its advocates and detractors have tended to give the impression that listeners effectively act as a multi-channel recording system permanently switched on through which the entirety of experience is channelled, indexed and stored away in memory. (Docherty and Foulkes 2014: 51)

Is it theoretically productive to assume that the exemplar store models what goes on in people’s brains? To answer this question, it is instructive to turn to an analogous debate in the neurophilosophy of visual perception. There, the assumption that perceptual experiences map onto the brain in a (coarse, but) direct fashion has long been known by the name of “homunculus fallacy”. More specifically, it has been claimed that the view that incoming visual information is directly projected into the brain as if onto a screen presupposes an imaginary little man, a “homunculus”, who sits in one’s brain, re-perceives the input, and interprets it “for” the subject. However, as Blakeslee and Blakeslee caution, the homunculus assumption

utterly fails to explain perception, understanding, and action: How does the homunculus perceive, understand, and act? The only way to ‘explain’ the homunculus’s abilities is to posit another, smaller homunculus inside ‘him.’ But then the same problem pops up, and you’re left with an endless series of Russian doll homunculi. [...] A philosopher (or a neuroscientist) commits the homunculus fallacy whenever he ‘explains’ something important about how the mind works by sidestepping the real difficulties of the problem and shifting them to another, unspecified level of explanation – where it remains just as mysterious as ever. (Blakeslee and Blakeslee 2008: 18)

In the case of exemplar-based models of language, the homunculus would be an instance that constantly monitors the exemplar store, selects the memory traces relevant to the task at hand, and performs on-line statistical analyses over them – and the same epistemological concerns as in vision would apply, i. e., we would not learn anything new about how the brain actually performs these tasks (cf. also Dennett 1992, 2006).

At first sight, it might seem that placing a filter between perception and storage might provide a way to escape the homunculus fallacy. This filter could, for instance, place exemplars that are more similar to each other along some dimension(s) in greater proximity to each other in the memory space, as is often assumed in exemplar-based models (cf. Gries 2012). But, in the absence of any stipulation as to how this could emerge bottom-up from networks of interacting neurons, such a procedure again presupposes some kind of centralized instance which selects relevant dimensions, compares and categorizes incoming tokens with respect to them, and constantly reorganizes the memory space, i. e., a kind of homunculus.

Against this background, I would like to suggest that it is not theoretically productive to assume that the exemplar store provides direct insights into the brain. On the other hand, it does, obviously, still provide insights into cognition, since it allows us to account for existing behavioral data (cf. Section 1). To define the cognitive modeling contribution of exemplar-based models in positive and more specific terms, it is again helpful to turn to a similar debate in another domain of cognitive science.

According to a taxonomy proposed by Marr in 1982 (Marr et al. 2010: 103–111), which is widely accepted in the (non-linguistic) probabilistic cognition community, the human mind can be modeled at three mutually constraining levels:

- (1) the *computational level*, which characterizes the nature of a cognitive problem and its ideal solution in abstract terms – this is the level at which statistical models of cognition are cast;
- (2) the *algorithmic level*, which identifies the cognitive representations and processes used to solve the problem at the computational level – this is the level at which cognitive psychology models operate;

- (3) the *implementational level*, which specifies the neural processes underlying the algorithmic level – this is the level of analysis of the cognitive neurosciences.

Importantly, it is usually assumed that adopting a probabilistic model at the computational level is noncommittal as to whether probabilities are explicitly represented at the underlying psychological and neural levels (Griffiths et al. 2010: 362). With this in mind, I would like to suggest that corpus-based cognitive linguistic models are cast at the computational level, i.e., at a level quite removed from the brain, and that they are cognitively realistic in the loose sense of imposing constraints on the underlying psychological and neural levels (Griffiths et al. 2012).

3 What is the added value of neurolinguistics to corpus-based Cognitive Linguistics?

The first reason why I consider it fruitful for corpus-based cognitive linguists to engage with neurolinguistics follows immediately from the preceding section: Research bridging the gap between levels of analysis broadens the scope of usage-based theorizing and promises significant progress towards increased cognitive realism (Griffiths et al. 2012: 263).

I consider this to be an end in itself, as well as a major motivating factor in the search for behavioral and brain correlates to theoretical assumptions. In my view, recent years of corpus-driven cognitive linguistics research have seen encouraging theoretical and experimental progress in relating the computational to the algorithmic level, for example by relating frequency metrics to cognitive psychology notions such as attention and automatization (cf. Gries and Ellis 2015; Küchenhoff and Schmid 2015; and the references therein), and by developing modeling techniques for statistical data based on principles of human learning (Milin et al. this issue).

Likewise, several recent neuroimaging studies have started to explore correlations between different kinds of corpus-derived frequency metrics and lexical as well as morphological processing, on the basis of different kinds of techniques (Blumenthal-Dramé et al. under revision; Frank et al. 2015; Fruchter and Marantz 2015; Graves et al. 2009; Hanna and Pulvermüller 2014; Hauk et al. 2008; Lewis et al. 2011; Linzen et al. 2013; Solomyak and Marantz 2009; Tremblay and Baayen 2010; Willems et al. 2016; Wilson et al. 2009). Moreover,

research on the neuro-functional underpinnings of the processing of more versus less predictable linguistic features at different levels – from sub-phonemic up to discourse-pragmatic – has attracted increasing scholarly interest over the last years, even if most of these studies do not (yet) operationalize predictability in terms of corpus-derived statistics (for reviews, cf. Huettig and Mani 2016; Kuperberg and Jaeger 2016; Lewis and Bastiaansen 2015; Van Petten and Luka 2012). All these studies show that neurolinguistic experiments make it possible to avoid the circularity of the homunculus fallacy by complementing statements cast at the computational level with converging evidence from a qualitatively distinct dimension.

However, I still understand researchers who are dismissive of hunting for brain correlates to theoretical linguistic constructs. One reason is that identifying systematic correlations between epistemological levels does not necessarily advance our understanding of the level that linguists are foremost interested in, the computational level. Second, the fact that an experimental result is in line with usage-based predictions does not entail that it is incompatible with predictions from competing theories – in other words, such a result does not necessarily tell us that the usage-based model is cognitively more realistic than competing models cast at the same level. Third, the more we expand our experimental repertoire, the more likely it becomes to find our predictions confirmed in at least one experiment tracking one specific processing stage in one specific brain region. However, we have no *a priori* criterion to determine what constitutes sufficient evidence in support of the cognitive realism of a claim.

I would therefore like to suggest an even more insightful way in which neurolinguistic experiments can contribute to establishing the cognitive realism of linguistic claims: By pitting the cognitive predictions of competing linguistic models against each other (Tressoldi et al. 2012). For example, Blumenthal-Dramé et al. (under revision) report an fMRI study that aims to adjudicate between cognitive predictions derived from usage-based theory, on the one hand, and Distributed Morphology (Halle and Marantz 1994), on the other hand. Distributed Morphology and related morphological processing accounts (e.g., Fruchter and Marantz 2015) assume that morphemes are cognitively realistic in the sense of being necessarily accessed in the first stages of the processing of morphologically complex words (e.g., *government*, *harmless*). Importantly, early access to morphemic constituents is assumed not to be modulated by higher-level properties like whole-word frequency. By contrast, a strong version of usage-based assumptions about frequency-induced chunk status predicts that whole-word properties should play a role from the first processing stages onwards.

Blumenthal-Dramé et al. demonstrate that early access to morphemes is modulated by log-transformed relative frequency (a measure which results from dividing whole-word by base frequency, cf. Hay 2001). Among other things, this metric is shown to correlate with activation in brain regions associated with competition between holistically represented entries (both simple and complex). Blumenthal-Dramé et al. conclude that knowledge about the usage of whole words influences early stages of processing, suggesting that usage-based assumptions concerning early morphological processing stages are more cognitively realistic than assumptions derived from Distributed Morphology.

Although this shows that neurolinguistic experiments have the potential to yield qualitatively new insights and to feed back into usage-based theorizing, I still see a major stumbling block in the way of cross-fertilization between Cognitive Linguistics and neurolinguistics: unrealistic expectations of what neurolinguistics can and cannot do for linguistic theorizing. The next section will therefore set straight a number of misconceptions that I frequently encounter.

4 Stumbling blocks in the way of cross-fertilization between corpus-based statistical models and neurolinguistics

Caricaturing a little bit, there are broadly three kinds of reactions that I get when telling other linguists that I do neurolinguistic experiments in the hope of contributing to linguistic theory:

- 1.) “Are you sure this is going to tell you anything? Isn’t there a brain correlate to everything anyways?”
- 2.) “Why don’t you just conduct behavioral experiments? Wouldn’t they tell you the same, but take less effort, time and money?”
- 3.) “Wow, that’s great, could you help me devise a neuroimaging experiment that tells me whether the difference between phenomenon A (e. g., derivational morphology) and phenomenon B (e. g., inflectional morphology) is categorical or gradient?”

In my view, all of these questions ultimately result from a widespread, but rather naive expectation that a “glimpse into people’s brain” can or should afford clear, simple and objective answers, whereas a look at their behavior can only yield a fuzzy and distorted picture – a view which I find reminiscent of the performance-

competence distinction. However, there is no reason to assume that neuroimaging data are clearer or less open to debate and interpretation than any other language-related data (corpus data, reaction time data, typological data, etc.).

Turning to the first question, there is, of course, a neural activation correlate to every instance of language use, and on top of that, this correlate will look different in the brains of different speakers, it will be modulated by a host of contextual and speaker-related variables, it will change over lifetime and, moreover, it will depend on the experimental design of your study, the spatial and temporal resolution of your neuroimaging technique, the processing stage under consideration, and the statistical method used in the analysis. But this only comes as a disappointment if you presuppose that brain data should be simpler than other kinds of language-related data – after all, correlations per se have never been an obstacle to linguistic theorizing (for example, every language form comes with a function, but this does not prevent linguists from doing research on language forms).

Second, compared to purely behavioral data, neuroimaging data add a qualitatively new dimension to our picture of the human mind. This becomes obvious when considering the fact that behavioral experiments compress the results of neural processes that are massively parallel, interactive and distributed in time and space into a single behavioral output (e. g., a button press), and that different neuro-functional processes can give rise to similar behavioral outputs. As a result, corpus-derived probabilistic measures can exhibit significant neural activation correlates in the absence of any behavioral correlates. For example, an MEG study by Linzen et al. (2013) shows that the entropy of a verb's subcategorization frame (SCF) distribution, which captures the degree of uncertainty about upcoming structural patterns, does not affect lexical decision times to verbs. By contrast, SCF distribution does correlate with activity in the anterior temporal lobe in a time window 200–300 ms after stimulus presentation.

Third, neuroimaging data leave (at least) as much room for different theoretical interpretations as any other language-related data, and they also face similar categorization problems: for example, how to group data points, how to delimit overlapping data clusters in space and time, and where to set significance thresholds. In other words, contrary to commonly held belief, categories are no more “given”, objective and discrete in the brain than they are elsewhere. As a result, I do not expect neuroimaging experiments on their own to be able to settle longstanding categorization issues such as whether generalizations about usage (e. g., probabilities) are intrinsic to grammar or extra-grammatical (Gahl and Garnsey 2004, 2006; Newmeyer 2003, 2006), or whether the difference between lexicon and grammar is a categorical or gradient one (Fedorenko et al. 2012). While neurolinguistic experiments can, of course,

add new data to the debate, these data won't take us very far as long as researchers do not agree on the theoretical question of what would actually constitute empirical evidence for or against a "cognitive gap" in the brain.

5 Conclusion

To summarize, in my view, neurolinguistics clearly has the potential to enhance the cognitive realism of theoretical linguistic claims. I therefore expect theoretical linguistics to increasingly draw upon neurolinguistic methods and findings. It is, however, necessary to stress that my notion of cognitive realism is not an absolute one: As outlined above, I consider it more fruitful to view cognitive realism as an ideal that can be approached through testing and comparing the predictive power of competing linguistic theories which purport to be cognitively realistic.

Moreover, I also hope to have shown that "looking into the brain" will not reduce or diminish anything – neither the complexity of theoretical debates, nor the number of modeling levels required, nor the role of the theoretical linguist. On the contrary, my prediction is that closer collaboration with neurolinguistics will rekindle discussions that have been going on in corpus-based Cognitive Linguistics for a long time (e. g., how to maximally avoid a priori theoretical commitments in empirical research, how to deal with converging and diverging evidence, or how to integrate inter- and intra-individual differences into theoretical models). In addition, this move will give rise to new theoretical challenges, such as the following: What follows for a theory that turns out to be less cognitively realistic than a competing one? How do cognitive linguistic predictions interface with existing neurological models? How can usage-based concepts be operationalized at the implementational level? For all these reasons, I do not expect neurolinguistics to ultimately absorb theoretical linguistics. In particular, I do not expect corpus-based Cognitive Linguistics to lose its *raison d'être* as a distinct strand of research that builds bridges between usage data and other levels of analysis.

References

- Amenta, S. & D. Crepaldi. 2012. Morphological processing as we know it: An analytical review of morphological effects in visual word identification. *Frontiers in Psychology* 3. 232. doi: 10.3389/fpsyg.2012.00232
- Arai, M. & F. Keller. 2013. The use of verb-specific information for prediction in sentence processing. *Language and Cognitive Processes* 28(4). 525–560.

- Arnon, I. & U. Cohen Priva. 2013. More than words: The effect of multi-word frequency and constituency on phonetic duration. *Language and Speech* 56(3). 349–371.
- Arnon, I. & N. Snider. 2010. More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language* 62(1). 67–82.
- Blakeslee, S. & M. Blakeslee. 2008. *The body has a mind of its own: How body maps in your brain help you do (almost) everything better*. New York: Random House Trade Paperbacks.
- Blumenthal-Dramé, A. 2012. *Entrenchment in usage-based theories: What corpus data do and do not reveal about the mind*. Berlin: Walter de Gruyter.
- Blumenthal-Dramé, A., V. Glauche, T. Bormann, C. Weiller, M. Musso & B. Kortmann. Under revision. Frequency and chunking in derived words: A parametric fMRI study. *Journal of Cognitive Neuroscience*.
- Bresnan, J. 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, J. & M. Ford. 2010. Predicting syntax: Processing dative constructions in American and Australian varieties of English. *Language* 86(1). 168–213.
- Bresnan, J. & J. Hay. 2008. Gradient grammar: An effect of animacy on the syntax of give in New Zealand and American English. *Lingua* 118(2). 245–259.
- Bybee, J. 2010. *Language, usage and cognition*. Cambridge, UK: Cambridge University Press.
- Chater, N. & M. Oaksford. 2008. *The probabilistic mind: Prospects for Bayesian cognitive science*. Oxford: Oxford University Press.
- Chater, N., J. B. Tenenbaum & A. Yuille. 2006. Probabilistic models of cognition: Conceptual foundations. *Trends in Cognitive Sciences* 10(7). 287–291.
- Dąbrowska, Ewa. 2016. Cognitive linguistics' seven deadly sins. *Cognitive Linguistics* 27(4).
- Demberg, V. & F. Keller. 2008. Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition* 109(2). 193–210.
- Dennett, D. C. 1992. *Consciousness explained*. Boston: Back Bay Books.
- Dennett, D. C. 2006. *Sweet dreams: Philosophical obstacles to a science of consciousness*. Cambridge, MA: Mit University Press Group Ltd.
- Divjak, D. 2016. The role of lexical frequency in the acceptability of syntactic variants: Evidence from *that*-clauses in Polish. *Cognitive Science*. doi: 10.1111/cogs.12335.
- Divjak, D. & S. Th. Gries. 2012. *Frequency effects in language representation – Vol. 2* (Trends in Linguistics. Studies and Monographs [TiLSM] 244.2). Berlin: Walter de Gruyter.
- Docherty, G. J. & P. Foulkes. 2014. An evaluation of usage-based approaches to the modeling of sociophonetic variability. *Lingua* 142. 42–56.
- Fedorenko, E., A. Nieto-Castañón & N. Kanwisher. 2012. Lexical and syntactic representations in the brain: An fMRI investigation with multi-voxel pattern analyses. *Neuropsychologia* 50(4). 499–513.
- Feldman, L. B., P. Milin, K. W. Cho, F. Moscoso del Prado Martín & P. A. O'Connor. 2015. Must analysis of meaning follow analysis of form? A time course analysis. *Frontiers in Human Neuroscience* 9. 111. doi: 10.3389/fnhum.2015.00111
- Fine, A. B., T. F. Jaeger, T. A. Farmer & T. Qian. 2013. Rapid expectation adaptation during syntactic comprehension. *PLOS ONE* 8(10). e77661.
- Frank, S. L., L. J. Otten, G. Galli & G. Vigliocco. 2015. The ERP response to the amount of information conveyed by words in sentences. *Brain and Language* 140. 1–11.

- Frisson, S., K. Rayner & M. J. Pickering. 2005. Effects of contextual predictability and transitional probability on eye movements during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31(5). 862–877.
- Fruchter, J. & A. Marantz. 2015. Decomposition, lookup, and recombination: MEG evidence for the full decomposition model of complex visual word recognition. *Brain and Language* 143. 81–96.
- Gahl, S. & S. M. Garnsey. 2004. Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language* 80(4). 748–775.
- Gahl, S. & S. M. Garnsey. 2006. Knowledge of grammar includes knowledge of syntactic probabilities. *Language* 82(2). 405–410.
- Gahl, S. & A. C. L. Yu. 2006. Introduction to the special issue on exemplar-based models in linguistics. *The Linguistic Review* 23(3). 213–216.
- Goldberg, A. E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
- Graves, W. W., R. Desai, C. Humphries, M. S. Seidenberg & J. R. Binder. 2009. Neural systems for reading aloud: A multiparametric approach. *Cerebral Cortex* 20(8). 1799–1815.
- Gries, S. Th. 2012. Corpus linguistics, theoretical linguistics, and cognitive/psycholinguistics: Towards more and more fruitful exchanges. *Language and Computers* 75(1). 41–63.
- Gries, S. Th. & D. Divjak. 2012. *Frequency effects in language learning and processing*. Berlin: Walter de Gruyter.
- Gries, S. Th. & N. C. Ellis. 2015. Statistical measures for usage-based linguistics. *Language Learning* 65(S1). 228–255.
- Gries, S. Th., B. Hampe & D. Schönefeld. 2005. Converging evidence: Bringing together experimental and corpus data on the association of verbs and constructions. *Cognitive Linguistics* 16(4). 635–676.
- Gries, S. Th., B. Hampe & D. Schönefeld. 2010. Converging evidence II: More on the association of verbs and constructions. In Sally Rice & John Newman (eds.), *Empirical and experimental methods in cognitive/functional research*, 59–72. Stanford, CA: CSLI Publications.
- Griffiths, T. L., N. Chater, C. Kemp, A. Perfors & J. B. Tenenbaum. 2010. Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences* 14(8). 357–364.
- Griffiths, T. L., M. Steyvers & J. B. Tenenbaum. 2007. Topics in semantic representation. *Psychological Review* 114(2). 211–244.
- Griffiths, T. & J. Tenenbaum. 2006. Statistics and the Bayesian mind. *Significance* 3(3). 130–133.
- Griffiths, T. L., E. Vul & A. N. Sanborn. 2012. Bridging levels of analysis for probabilistic models of cognition. *Current Directions in Psychological Science* 21(4). 263–268.
- Halle, M. & A. Marantz. 1994. Some key features of distributed morphology. *MIT Working Papers in Linguistics* 21. 275–288.
- Hanna, J. & F. Pulvermüller. 2014. Neurophysiological evidence for whole form retrieval of complex derived words: A mismatch negativity study. *Frontiers in Human Neuroscience* 8. 886.
- Hare, M., M. K. Tanenhaus & K. McRae. 2007. Understanding and producing the reduced relative construction: Evidence from ratings, editing and corpora. *Journal of Memory and Language* 56(3). 410–435.

- Hauk, O., M. H. Davis & F. Pulvermüller. 2008. Modulation of brain activity by multiple lexical and word form variables in visual word recognition: A parametric fMRI study. *NeuroImage* 42(3). 1185–1195.
- Hay, J. 2001. Lexical frequency in morphology: Is everything relative? *Linguistics* 39(6). 1041–1070.
- Hay, J. & R. H. Baayen. 2005. Shifting paradigms: Gradient structure in morphology. *Trends in Cognitive Sciences* 9(7). 342–348.
- Huetig, F. & N. Mani. 2016. Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience* 31(1). 19–31.
- Jaeger, T. F. & N. E. Snider. 2013. Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition* 127(1). 57–83.
- Jones, M. N. & D. J. K. Mewhort. 2007. Representing word meaning and order information in a composite holographic lexicon. *Psychological Review* 114(1). 1–37.
- Kamide, Y. 2012. Learning individual talkers' structural preferences. *Cognition* 124(1). 66–71.
- Kapatsinski, V. & J. Radicke. 2009. Frequency and the emergence of prefabs: Evidence from monitoring. In R. Corrigan, E. A. Moravcsik, H. Ouali & K. Wheatley (eds.), *Formulaic language. Vol. 2: Acquisition, loss, psychological reality, and functional explanations*, 499–520. Amsterdam: Benjamins.
- Kleinschmidt, D. F. & T. F. Jaeger. 2015. Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review* 122(2). 148–203.
- Küchenhoff, H. & H.-J. Schmid. 2015. Reply to "More (old and new) misunderstandings of collocation analysis: On Schmid & Küchenhoff" by Stefan Th. Gries. *Cognitive Linguistics* 26(3). 537–547.
- Kuperberg, G. R. & T. F. Jaeger. 2016. What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience* 31(1). 32–59.
- Levy, R. 2008. Expectation-based syntactic comprehension. *Cognition* 106(3). 1126–1177.
- Lewis, A. G. & M. Bastiaansen. 2015. A predictive coding framework for rapid neural dynamics during sentence-level language comprehension. *Cortex* 68. 155–168.
- Lewis, G., O. Solomyak & A. Marantz. 2011. The neural basis of obligatory decomposition of suffixed words. *Brain and Language* 118(3). 118–127.
- Linzen, T. & T. F. Jaeger. 2016. Uncertainty and expectation in sentence processing: Evidence from subcategorization distributions. *Cognitive Science* 40(6). 1287–1585.
- Linzen, T., A. Marantz & L. Pykkänen. 2013. Syntactic context effects in visual word recognition: An MEG study. *The Mental Lexicon* 8(2). 117–139.
- Marr, D., S. Ullman, & T. A. Poggio. 2010. *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge, MA: MIT Press.
- McDonald, S. A. & R. C. Shillcock. 2003. Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research* 43(16). 1735–1751.
- Milin, P., D. Divjak, S. Dimitrijević & R. H. Baayen. 2016. Towards cognitively plausible data science in language research. *Cognitive Linguistics* 27(4).
- Newmeyer, F. J. 2003. Grammar is grammar and usage is usage. *Language* 79(4). 682–707.
- Newmeyer, F. J. 2006. On Gahl and Garnsey on grammar and usage. *Language* 82(2). 399–404.
- Perfors, A., J. B. Tenenbaum, T. L. Griffiths & F. Xu. 2011. A tutorial introduction to Bayesian models of cognitive development. *Cognition* 120(3). 302–321.

- Schmidtke, D., V. Kuperman, C. L. Gagné & T. L. Spalding. 2015. Competition between conceptual relations affects compound recognition: The role of entropy. *Psychonomic Bulletin and Review* 23(2). 556–570.
- Siyanova-Chanturia, A., K. Conklin & W. J. B. Van Heuven. 2011. Seeing a phrase “time and again” matters: The role of phrasal frequency in the processing of multiword sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37(3). 776–784.
- Smith, N. J. & R. Levy. 2013. The effect of word predictability on reading time is logarithmic. *Cognition* 128(3). 302–319.
- Snider, N. & I. Arnon. 2012. A unified lexicon and grammar? Compositional and noncompositional phrases in the lexicon. In S. Gries & D. Divjak (eds.), *Frequency effects in language*, 127–163. Berlin: Mouton de Gruyter.
- Solomyak, O. & A. Marantz. 2009. Evidence for early morphological decomposition in visual word recognition. *Journal of Cognitive Neuroscience* 22(9). 2042–2057.
- Tenenbaum, J. B., C. Kemp, T. L. Griffiths & N. D. Goodman. 2011. How to grow a mind: Statistics, structure, and abstraction. *Science* 331(6022). 1279–1285.
- Tily, H., S. Gahl, I. Arnon, N. Snider, A. Kothari & J. Bresnan. 2009. Syntactic probabilities affect pronunciation variation in spontaneous speech. *Language and Cognition* 1(2). 147–165.
- Tremblay, A. & R. H. Baayen. 2010. Holistic processing of regular four-word sequences: A behavioral and ERP study of the effects of structure, frequency, and probability on immediate free recall. In D. Wood (ed.), *Perspectives on formulaic language: Acquisition and communication*, 151–173. London: Continuum.
- Tremblay, A. & B. V. Tucker. 2011. The effects of N-gram probabilistic measures on the recognition and production of four-word sequences. *The Mental Lexicon* 6(2). 302–324.
- Tremblay, A., B. Derwing, G. Libben & C. Westbury. 2011. Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks: Lexical bundle processing. *Language Learning* 61(2). 569–613.
- Tressoldi, P. E., F. Sella, M. Coltheart & C. Umiltà. 2012. Using functional neuroimaging to test theories of cognition: A selective survey of studies from 2007 to 2011 as a contribution to the Decade of the Mind Initiative. *Cortex* 48(9). 1247–1250.
- Van Petten, C. & B. J. Luka. 2012. Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology* 83(2). 176–190.
- Willems, R. M., S. L. Frank, A. D. Nijhof, P. Hagoort & A. van den Bosch. 2016. Prediction during natural language comprehension. *Cerebral Cortex* 26(6). 2506–2516.
- Wilson, M. P. & S. M. Garnsey. 2009. Making simple sentences hard: Verb bias effects in simple direct object sentences. *Journal of Memory and Language* 60(3). 368–392.
- Wilson, S. M., A. L. Isenberg, & G. Hickok 2009. Neural correlates of word production stages delineated by parametric modulation of psycholinguistic variables. *Human Brain Mapping* 30(11). 3596–3608.