

# Language and Speech

<http://las.sagepub.com/>

---

## Introduction to the Special Issue: Parsimony and Redundancy in Models of Language

Daniel Wiechmann, Elma Kerz, Neal Snider and T. Florian Jaeger  
*Language and Speech* published online 28 June 2013  
DOI: 10.1177/0023830913490877

The online version of this article can be found at:  
<http://las.sagepub.com/content/early/2013/06/25/0023830913490877.citation>

---

Published by:



<http://www.sagepublications.com>

**Additional services and information for *Language and Speech* can be found at:**

**Email Alerts:** <http://las.sagepub.com/cgi/alerts>

**Subscriptions:** <http://las.sagepub.com/subscriptions>

**Reprints:** <http://www.sagepub.com/journalsReprints.nav>

**Permissions:** <http://www.sagepub.com/journalsPermissions.nav>

>> [OnlineFirst Version of Record](#) - Jun 28, 2013

[What is This?](#)

# Introduction to the Special Issue: Parsimony and Redundancy in Models of Language

Language and Speech

0(0) 1–8

© The Author(s) 2013

Reprints and permissions:

sagepub.co.uk/journalsPermissions.nav

DOI: 10.1177/0023830913490877

las.sagepub.com

**Daniel Wiechmann**

University of Amsterdam, The Netherlands

**Elma Kerz**

RWTH Aachen University, Germany

**Neal Snider**

Nuance Communications, Inc., USA

**T. Florian Jaeger**

University of Rochester, USA

One of the most fundamental goals in linguistic theory is to understand the nature of linguistic knowledge, that is, the representations and mechanisms that figure in a cognitively plausible model of human language-processing. The past 50 years have witnessed the development and refinement of various theories about what kind of ‘stuff’ human knowledge of language consists of, and technological advances now permit the development of increasingly sophisticated computational models implementing key assumptions of different theories from both rationalist and empiricist perspectives. The present special issue does not aim to present or discuss the arguments for and against the two epistemological stances or discuss evidence that supports either of them (cf. Bod, Hay, & Jannedy, 2003; Christiansen & Chater, 2008; Hauser, Chomsky, & Fitch, 2002; Oaksford & Chater, 2007; O’Donnell, Hauser, & Fitch, 2005). Rather, the research presented in this issue, which we label *usage-based* here, conceives of linguistic knowledge as being induced from experience. According to the strongest of such accounts, the acquisition and processing of language can be explained with reference to general cognitive mechanisms alone (rather than with reference to innate language-specific mechanisms). Defined in these terms, usage-based approaches encompass approaches referred to as *experience-based*, *performance-based* and/or *emergentist* approaches (Arnon & Snider, 2010; Bannard, Lieven, & Tomasello, 2009; Bannard & Matthews, 2008; Chater & Manning, 2006; Clark & Lappin, 2010; Gerken, Wilson, & Lewis, 2005; Gomez, 2002;

**Corresponding author:**

Daniel Wiechmann, RWTH Aachen University, Department of English, American and Romance Language Studies, Kármánstr, 17/19, D-52062 Aachen, Germany.

Email: wiechmann@anglistik.rwth-aachen.de

MacWhinney, 1998, 2005; Marcus, Vijayan, Bandi Rao, & Vishton, 1999; O'Grady, 2010; Elman et al. (1996); Tabor & Tanenhaus, 1999; Chang, Dell, & Bock, 2006). This special issue sets out to address fundamental questions regarding parsimony of representation and mechanism that arise when such a standpoint is taken.

With virtually all theories of language, usage-based approaches assume that the vast expressive power of natural languages, which allows speakers to produce an uncountable set of possible utterances, resides in their cognitive capacity to store and combine units. However, representational nativist and usage-based accounts offer radically different views on how the underlying knowledge should be conceived of. Representational nativist accounts (Chomsky, 1981; Pinker, 1995), or at least those that are explicit about learning mechanisms (cf. Pollard & Sag, 1994), argue that linguistic knowledge comprises categorically different types of entities, most notably lexical knowledge and grammatical principles or rules. In this view, these ingredients of language are acquired in qualitatively different ways in language development and are each regulated by their own set of mechanisms in language processing (see also some of the dominant psycholinguistic accounts, in particular, of language production, Bock & Levelt, 2002; Levelt, 1993; Pickering & Garrod, 2004). Symbolic units are assumed to be learned inductively. Grammatical rules, on the other hand, are not learned but acquired through some process of parameterization of an innately specified system of possibilities (cf. Universal Grammar, Chomsky, 1981; Lenneberg, 1967; Pinker, 1995).

Usage-based approaches offer a different view on the building blocks of language. A common assumption in usage-based accounts is that *all* linguistic knowledge is learned based on a single set of acquisition processes. All established units in the system are the product of some learning algorithm that analyzes linguistic utterances into constitutive fragments (sometimes called *chunks*). Grammatical rules then are not another type of stuff but higher-level abstractions of experience (stored generalizations) (cf. Beekhuizen, Bod, & Zuidema, this issue; Goldberg, 2006; Post & Gildea, this issue). In production, structures at different levels of abstraction can be combined into more complex structures allowing speakers to potentially express an uncountable set of ideas. Since the human capacity to combine 'cognitive units', like the capacity to abstract, is part and parcel of a set of general cognitive mechanisms also employed outside of language, these accounts leave open the possibility that no additional set of language-specific mechanisms is required.

During language development, the extension of the set of mentally represented linguistic units seems to be subject to change (above and beyond regular vocabulary growth): previously unanalyzed chunks are further analyzed, creating new building blocks (Bybee, 2010 for overviews; Perfors, Tanenbaum, & Regier, (2011); Tomasello, 2003). There is also some emerging evidence that similar learning processes might continue throughout life (Fine, Jaeger, Qian, & Farmer, submitted; Jaeger & Snider, 2013; Kamide, 2012). For example, frequent exposure to structures can change the way they are being produced or comprehended, and there is some evidence that these changes can persist (Kaschak, Kutta, & Schatschneider, 2011). Such changes can also involve the acquisition of novel syntactic structures in adults (e.g., from previously unencountered dialects, Kaschak & Glenberg, 2004).

Furthermore, it has been proposed that the repeated use of strings of units may lead to the establishment of new, complex (fully compositional) units, which can be retrieved from memory (due to general cognitive processes such as routinization, also known as automatization or entrenchment; Langacker, 1987, 2008).

The fact that that the same sentence can, in principle, be produced through the combination of different subtrees of varying size and degree of abstraction leads to a key issue, namely what set of units is retained in an adult system, and to what extent is the adult system characterized by representational (= storage) redundancy. The spectrum of conceivable (and actualized) theoretical

positions includes anything from theories that envision a maximally parsimonious representation of linguistic knowledge to theories that postulate massive representational redundancy. If we are willing to assume that there is some level of representational redundancy in the system, that is, if fragments of different size and degrees of abstraction may coexist, a central question to be answered is: How can a given level of representational redundancy be justified? Which structures should be stored and why? Is the principle of simplicity according to which the cognitive system should prefer the simplest possible solution the correct guiding principle? And if so, what are competing aspects of simplicity (or *utility*)? For example, if representational redundancy facilitates processing – because the existence of chunks allows a faster mapping of form to meaning (compared to the computation of compositions) – how do language users strike a balance between these two pressures? In light of these considerations, this special issue is also geared to discuss the idea that an elegant model of linguistic knowledge can be one that exhibits high degrees of parsimony at the level of processing and/or at the level of representation. Recent research in the computational and mathematical modeling of language learning and processing has made substantial progress in the attempts to answer these questions (Albright & Hayes, 2003; Alishahi & Stevenson, 2008; Daelemans & Van den Bosch, 2005; Griffiths, Chater, Kemp, Perfors, & Tenenbaum 2010; Gelman, 2004; Goldwater, Griffiths & Johnson, 2009; Griffiths & Tenenbaum, 2010; Li & Vitanyi, 2008, O'Donnell, 2011, for a review; Perfors, Tenenbaum, & Regier, 2011; Post & Gildea, 2009; Tenenbaum & Griffiths, 2001; Vallabha, McClelland, Pons, Werker, & Amano, 2007).

## **Computational incarnations of usage based theory: Example-driven stochastic versus exemplar-based models**

The fundamental principles of usage-based approaches can be instantiated in various types of computational architectures at different levels of analysis. What is more, some aspects of usage-based theory are, at this point, more appropriately addressed with some architectures, while other aspects are currently better handled by other architectures. For example, approaches that seek to emphasize the highly distributed nature of linguistic knowledge and issues of emergence lean towards subsymbolic, connectionist architectures. Such approaches also exhibit a stronger commitment to what has been termed *biological realism* (e.g., Koch & Segev, 1989) and often explore issues of how linguistic knowledge could plausibly be realized in brain-like systems. They thus target a level of description closer to what Marr (1982) termed the implementational level (cf. Marr, 1982; for a more comprehensive discussion, cf. Christiansen & Chater, 2008; Elman, 2011; McClelland et al., 2010; Oaksford & Chater, 2007; Rumelhart, McClelland, & the PDP Research Group, 1986). The stronger commitment to biological realism typically necessitates that current subsymbolic models of syntax use small fragments of grammar and small vocabularies raising the question of how well these models will scale up. The knowledge acquired by connectionist models is also often not transparent, and hence harder to interpret. This has been argued to make it more difficult to ascertain their fit against available experimental data (Albright & Hayes, 2003; Alishahi, 2011; Goldsmith, 2001). The models presented in this volume are situated at a higher level of description, that is, a level closer to what Marr termed the algorithmic level. It is at this higher-level that questions about representational units of the type targeted here, and by implication representational redundancy, are most naturally situated.

The approaches discussed in this issue run the gamut from what might be termed *example-driven stochastic* models, in which more abstract and more concrete generalizations co-exist, to *radical exemplar-based* models in which all and only exemplars are stored and generalization is

achieved through analogical reasoning. The latter is represented by *Memory-based Language Processing* (MBLP, van den Bosch & Daelemans, this issue). The former class is represented here by *tree-substitution grammars* (TSG, Post & Gildea, this issue) and *Data Oriented Parsing* models (DOP, Beekhuizen, Bod, & Zuidema, this issue) each equipped with a strategy to justify postulated levels of representational redundancy (TSG induction through Bayesian learning, DOP with Bayesian model merging; cf. O'Donnell, 2011 for an in-depth discussion). These models also differ with respect to other foundational assumptions, for example, assumptions about the structure of the stored units, that is, whether or not the units exhibit hierarchical structure, or the extent to which models depend on the fragments they have memorized. The MBLP approach, for example, employs analogical reasoning that frees it from the contingencies of the training data.

Despite their architectural differences, the models mentioned above agree that the information carried by (near-)compositional multi-word units is crucial for high-performance learning and processing and they capture this information by storing high numbers of examples (either represented as sets of features or tree-fragments). The latest variants of the models mentioned so far pursue different strategies that either target the reduction of the postulated instance space by way of finding the optimal number of representations (best subtree-approaches) or employ some kind of compression mechanism that speeds up processing.

A recent alternative computational architecture seeking to circumvent the problem of positing possibly hundreds of millions of representations has grown out of research into discrimination learning. The *Naïve Discriminative Reader* (NDR) model (cf. Baayen, 2011; Baayen, Hendrix, & Ramscar, this issue) represents a decompositional approach that can account for the observed effects typically interpreted as support for massive representation (e.g., the *n*-gram frequency effect discussed below) without storing exemplars and without explicit rule induction.

## 2 Experimental evidence

In pursuit of experimental evidence indicating the representational status of a linguistic form, that is, whether a form is holistically stored rather than being the product of some kind of combinatorial process, psycholinguists have proposed to employ behavioral measures, which in light of recent considerations has proved to be problematic. The general reasoning applied in this research (e.g., Alegre & Gordon, 1999; Baayen, Dijkstra, & Schreuder, 1997; Baayen, Schreuder, deJong, & Krott, 2002; Taft, 1979) can be illustrated on the basis of a classic study from the area of morphological processing, namely Taft (1979). Building on earlier experiments since Rosenberg, Coyle and Porter (1966), Taft investigated affixed words in English (e.g., noun + plural *-s*). The central idea underlying his experiments was to interpret frequency effects on reaction time differentials in lexical decision tasks, as a diagnostic for holistic (or analytic) storage: If whole-form frequency of a morphologically complex unit (e.g., *things*) affects processing independently of the frequency of the alleged components (*thing* + *s*), then this was interpreted as evidence for holistic storage. Motivated by usage-based theory, this argument from '*whole-form frequency* to *whole-form storage* has been transferred to the syntactic domain and questions about holistic storage of (even fully compositional) multi-word units: For example, Bannard and Matthews (2008) report that 2-3 year-old children process frequent multi-word phrases faster than infrequent ones and interpreted this as evidence for the storage of these phrases. Extending this study, Arnon and Snider (2010) found similar word *n*-gram frequency effects in adult language production. Revisiting the validity of the *whole-form frequency* to *whole-form representation* reasoning, Baayen and colleagues (Baayen, Hendrix & Ramscar, this issue; Baayen, Milin, Filipovic Durdevic, Hendrix, & Marelli, 2011) show that the phrase frequency effects reported

in Arnon and Snider (2010) can in fact be re-produced by a model that does not assume stored representations. In light of these considerations, the experimental research presented in the second part of this issue does not aim to provide direct behavioral evidence for or against particular representational assumptions but informs both theory-formation and computational modeling in a more indirect way. For example, Arnon and Cohen-Priven (this issue) combine experimental and corpus-based methodologies to investigate if effects of multi-word statistics on phonetic duration are modulated by higher order properties of the multi-word sequences, specifically their syntactic properties. Their results thus inform the further development of computational architectures with regard to structural assumptions. Holsinger (this issue) examines the role of structural information in the processing of idiomatic strings, which exhibit both word-like and structure-like properties. For example, while some idioms have idiosyncratic semantic properties suggesting that they are best conceived of as ‘big-words’, which need to be stored, not all of them are compositionally-opaque, allowing for different possibilities of representation. Furthermore, recent research has demonstrated that idioms can give rise to structural priming effects suggesting that they have internal structure (Konopka & Bock, 2009). On the basis of eye tracking experiments, Holsinger asks if syntactic information influences activation of idiomatic meaning during on-line comprehension and explores the implications of current models of idiom representation. Wedel, Jackson and Kaplan (this issue) investigate another assumption of usage based theories, namely that individual usage-events influence long-term language change, and by implication the shapes of individual representations of linguistic knowledge (Bybee, 2010). Wedel and colleagues investigate the conditions of phoneme merger, the loss of phonetic distinction between, for example, the vowels in the words *caught* and *cot*, which has been explained with reference to the amount of ‘work’ it does, or the *functional load* it bears, in information transmission (cf. Aylett & Turk, 2006; Jaeger, 2010; Piantadosi, Tily, & Gibson, 2011; Tily & Kuperman, 2012; for a recent overview, see Jaeger, & Tily, 2011). Based on statistical modeling of data from 19 systems of phoneme contrasts from nine languages, Wedel and colleagues pursue a refined version of the functional load hypothesis and investigate the relationship between phoneme-contrast merger probability and various variables. Among other things, they find evidence that that syntactic category and frequency relationships in minimal lemma pairs govern the loss of phoneme contrasts in language change, a finding that fits the predictions of usage-based models very well.

In conclusion, starting from the assumption that linguistic knowledge is induced from experience, this special issue explores a key problem in the usage-based framework, namely the balancing of storage and computation of units in language learning, comprehension and production. We recast this issue in terms of parsimony and redundancy in representation and processing. The papers in this special issue provide significant contributions to our understanding of the nature of linguistic knowledge by offering important insights from computational and experimental perspectives.

### Acknowledgements

We would like to take this opportunity to extend our gratitude to all people who have contributed to the success of this issue directly or indirectly. This special issue has evolved from a workshop of the same name at the 85th meeting of the Linguistic Society of America in 2011. We thank all contributors and the audience of that workshop for sharing their ideas. We thank the many expert reviewers of the proposed manuscripts for their highly informed and constructive criticism. We thank HLP Lab manager Andrew Watts for help with the organization of the workshop, Tim Bunnell and Irene Vogel for their interest in this work and, last but not least, Jim Polikoff for his outstanding editorial management and overall great support. Above all we would like to thank the contributors to this volume. Partial funding for this project came from NSF IIS-1150028 CAREER and an Alfred P. Sloan Fellowship to T. Florian Jaeger.

## References

- Albright, A., & Hayes, B. (2003). Rules vs. analogy in English past tenses: A computational/experimental study. *Cognition*, *90*, 119–161.
- Alegre, M., & Gordon, P. (1999). Frequency effects and the representational status of regular inflections. *Journal of Memory and Language*, *40*, 41–61.
- Alishahi, A. (2011). *Computational modeling of human language acquisition: Synthesis lectures on human language technologies*. San Rafael, CA: Morgan & Claypool.
- Alishahi, A., & Stevenson, S. (2008). A computational model of early argument structure acquisition. *Cognitive Science: A Multidisciplinary Journal*, *32*, 789–834.
- Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language*, *62*, 67–82.
- Aylett, M., & Turk, A. (2006). Language redundancy predicts syllabic duration and the spectral characteristics of vocal syllable nuclei. *Journal of the Acoustical Society of America*, *119*, 3048–3058.
- Bannard, C., Lieven, E., & Tomasello, M. (2009). Modeling children's early grammatical knowledge. *Proceedings of the National Academy of Sciences*, *106*, 17284–17289.
- Bannard, C., & Matthews, D. (2008). Stored word sequences in language learning. *Psychological Science*, *19*, 241–248.
- Baayen, R. H. (2011) Corpus linguistics and naive discriminative learning. *Brazilian Journal of Applied Linguistics*, *11*, 295–328.
- Baayen, R. H., Dijkstra, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, *37*, 94–117.
- Baayen, R. H., Milin, P., Filipovic Durdevic, D., Hendrix, P., & Marelli, M. (2011). An amorphous model for morphological processing in visual comprehension based on naive discriminative learning. *Psychological Review*, *118*, 438–482.
- Baayen, R. H., Schreuder, R., De Jong, N. H., & Krott, A. (2002) Dutch inflection: The rules that prove the exception. In S. Nootboom, F. Weerman & F. Wijnen (Eds.), *Storage and computation in the language faculty* (pp. 61–92). Dordrecht, The Netherlands: Kluwer.
- Bock, K., & Levelt, W. (2002). Language production: Grammatical encoding. *Psycholinguistics: Critical concepts in psychology*, *5*, 405–452.
- Bod, R., Hay, J., & Jannedy, S. (Eds.). (2003). *Probabilistic linguistics*. Cambridge, MA: The MIT Press.
- Bybee, J. (2010). *Language, usage and cognition*. Cambridge, UK: Cambridge University Press.
- Chang, F., Dell, G. S., & Bock, K. (2006). Becoming syntactic. *Psychological Review*, *113*(2), 234–272.
- Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences*, *10*, 335–344.
- Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*, 489–509.
- Chomsky, N. (1981). *Lectures on government and binding*. Berlin, Germany: Mouton de Gruyter.
- Clark, A., & Lappin, S. (2010). *Linguistic nativism and the poverty of the stimulus*. Oxford, UK: Wiley Blackwell.
- Daelemans, W., & Van den Bosch, A. (2005). *Memory-based language processing*. Cambridge, UK: Cambridge University Press.
- Elman, L. J., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., & Plunkett, K. (1996). *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge, MA: MIT Press.
- Elman, J. L. (2011). Lexical knowledge without a mental lexicon? *The Mental Lexicon*, *60*, 1–33.
- Fine, A. B., Jaeger, T. F., Qian, T., & Farmer, T. (submitted). Rapid expectation adaptation during syntactic comprehension. *Cognition*. Manuscript submitted for publication.
- Frank, M. C., Goldwater, S., Griffiths, T. L., & Tenenbaum, J. B. (2010). Modeling human performance in statistical word segmentation. *Cognition*, *117*, 107–125.
- Gelman, A. (2004). *Bayesian data analysis*. Boca Raton, FL: CRC press.
- Gerken, L., Wilson, R., & Lewis, W. (2005). Infants can use distributional cues to form syntactic categories. *Journal of Child Language*, *32*, 249–268.

- Goldberg, A. (2006). *Constructions at work: The nature of generalization in language*. Oxford, UK: Oxford University Press.
- Goldsmith, J. (2001). Unsupervised learning of the morphology of a natural language. *Computational linguistics*, 27, 153–198.
- Goldwater, S., Griffiths, T., & Johnson, M. (2009). A Bayesian framework for word segmentation: Exploring the effects of context. *Cognition* 112, 21–54.
- Gómez, R. (2002). Variability and detection of invariant structure. *Psychological Science*, 13, 431–436.
- Griffiths, T. L., Chater, N., Kemp, C., Perfors, A., & Tenenbaum, J. B. (2010). Probabilistic models of cognition: Exploring representations and inductive biases. *Trends in Cognitive Sciences*, 14, 357–364.
- Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298, 1569–1579.
- Jaeger, T. F. (2010). Redundancy and reduction: Speakers manage syntactic information density. *Cognitive psychology*, 61, 23–62.
- Jaeger, T. F., & Snider, N. (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. doi:10.1016/j.cognition.2012.10.013
- Jaeger, T. F., & Tily, H. (2011). On language 'utility': Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2, 323–335.
- Kaschak, M. P., & Glenberg, A. M. (2004). This construction needs learned. *Journal of experimental psychology. General*, 133, 450–467.
- Kaschak, M. P., Kutta, T. J., & Schatschneider, C. (2011). Long-term cumulative structural priming persists for (at least) one week. *Memory & Cognition*, 39, 381–388.
- Kamide, Y. (2012). Learning individual talkers' structural preferences. *Cognition*, 124, 66–71.
- Koch, C., & Segev, I. (1989). *Methods in neural modeling: From synapses to networks*. Cambridge, MA: MIT Press.
- Konopka, A. E., & Bock, J. K. (2009). Lexical or syntactic control of sentence formulation? Structural generalizations from idiom production. *Cognitive Psychology*, 58, 68–101.
- Langacker, R. W. (1987). *Foundations of cognitive grammar, vol. 1, theoretical prerequisites*. Stanford, CA: Stanford University Press.
- Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. New York, NY: Oxford University Press.
- Lenneberg, E. (1967). *Biological foundations of language*. New York, NY: John Wiley & Sons.
- Levelt, W. J. M. (1993). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Li, M., & Vitányi, P. (2008). *An introduction to kolmogorov complexity and its applications*. New York, NY: Springer-Verlag.
- MacWhinney, B. (1998). Models of the emergence of language. *Annual Review of Psychology*, 49, 199–227.
- MacWhinney, B. (2005). The emergence of grammar from perspective. In D. Pecher & R. Zwaan (Eds.), *Grounding cognition: The role of perception and action in memory, language and thinking* (pp. 198–223). Cambridge, UK: Cambridge University Press.
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283, 77–80.
- Marr, D. (1982). *Vision*. San Francisco, CA: W.H. Freeman.
- McClelland, J. L., Botvinick, M. M., Noelle, D. C., Plaut, D. C., Rogers, T.T., Seidenberg, M. S., & Smith, L. B. (2010). Letting structure emerge: Connectionist and dynamical systems approaches to understanding cognition. *Trends in Cognitive Sciences*, 14, 348–356.
- Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford, UK: Oxford University Press.
- O'Donnell, T. J. (2011). *Productivity and reuse in language* (Doctoral dissertation). Harvard University, Cambridge, MA.
- O'Donnell, T. J., Hauser, M. D., & Fitch, W. T. (2005). Using mathematical models of language experimentally. *Trends in Cognitive Sciences*, 9, 284–289.



- O'Grady, W. (2010). Emergentism. In P. Hogan (Ed.), *The Cambridge encyclopedia of the language sciences* (pp. 274–76). Cambridge, UK: Cambridge University Press.
- Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 3526–9.
- Perfors, A., Tenenbaum, J. B., & Regier, T. (2006). Poverty of the stimulus? A rational approach. In R. Sun, & N. Miyake (Eds.), *Proceedings of the 28th annual conference of the Cognitive Science Society* (pp. 663–668). Austin, TX: Cognitive Science Society.
- Perfors, A., Tenenbaum, J. B., & Regier, T. (2011). The learnability of abstract syntactic principles. *Cognition*, 306–338.
- Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27, 169–226.
- Pinker, S. (1995). *The language instinct: The new science of language and mind*. London, UK: Penguin.
- Pollard, C., & Sag, I. (1994). *Head-driven phrase structure grammar*. Chicago, IL: University of Chicago Press.
- Rosenberg, S., Coyle, P. J., & Porter, W. L. (1966). Recall of adverbs as a function of the frequency of their adjective roots. *Journal of Verbal Learning and Verbal Behavior*, 5, 65–76.
- Rumelhart, D. E., & McClelland, J. L., & the PDP Research Group. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.
- Post, M., & Gildea, D. (2009). Bayesian learning of a tree substitution grammar. In *Proceedings of the ACL-IJCNLP 2009 conference short papers* (pp. 45–48), Suntec, Singapore.
- Tabor, W., & Tannenhaus, M. K. (1999). Dynamical systems of sentence processing. *Cognitive Science*, 23, 491–515.
- Taft, M. (1979). Recognition of affixed words and the word frequency effect. *Memory & Cognition*, 7, 263–272.
- Tenenbaum, J., & Griffiths, T. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, 24, 629–640.
- Tily, H., & Kuperman, V. (2012). Rational phonological lengthening in spoken Dutch. *Journal of the Acoustical Society of America*, 132, 3935–3940.
- Tomaseello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.