



Topics in Cognitive Science 2 (2010) 367–373

Copyright © 2010 Cognitive Science Society, Inc. All rights reserved.

ISSN: 1756-8757 print / 1756-8765 online

DOI: 10.1111/j.1756-8765.2010.01102.x

Plus or Minus 30 Years in the Language Sciences

Elissa L. Newport

Department of Brain & Cognitive Sciences, University of Rochester

Received 1 November 2009; received in revised form 17 May 2010; accepted 17 May 2010

Abstract

The language sciences—Linguistics, Psycholinguistics, and Computational Linguistics—have not been broadly represented at the Cognitive Science Society meetings of the past 30 years, but they are an important part of the heart of cognitive science. This article discusses several major themes that have dominated the controversies and consensus in the study of language and suggests the most pressing issues of the future. These themes include differences among the language science disciplines in their view of *numbers and symbols* and of *modular and distributed cognition*, and the need for an increasing prominence of questions concerning *language and the brain*.

Keywords: Language sciences; Linguistics; Psycholinguistics; Statistical learning; Probabilistic approaches; Symbols and rules; Modularity; Language and the brain

1. Introduction

In this paper, I will be discussing the past and the future in the interdisciplinary language sciences: the set of disciplines that study language, including Linguistics, Psycholinguistics, and Computational Linguistics. This field is not typically the same as the audience that attends the Cognitive Science Society meetings, so the programs of the last 30 years of the Cognitive Science Society would not represent this part of cognitive science very well. Rather than addressing the past and future 30 years at the Cognitive Society meetings, then, I would like to discuss issues and dimensions of change in this field more broadly over the last 30 years, as well as those that might be prominent in the future. I hope that this will stimulate readers to think also about dimensions of focus and change they see in the past and those they think will characterize the future.

Correspondence should be sent to Elissa L. Newport, Department of Brain & Cognitive Sciences, Meliora Hall 314–River Campus, University of Rochester, Rochester, NY 14627. E-mail: newport@bcs.rochester.edu

The issues I will address are as follows: *numbers and symbols*; *modular and distributed cognition*; a very quick point on methodologies—*changes from print to real-time real-world language*; and I will end by discussing *language and the brain*.

2. Numbers and symbols

Let me start with *numbers and symbols*. The disciplines that study language have an extremely interesting difference in their first principles about qualitative versus quantitative, or symbolic versus numerical kinds of representations and processes, and this has led to a very interesting set of interactions and shifts—but also arguments and disagreements—in the field of language sciences.

Though there are certainly exceptions, I would characterize Formal Linguistics as, by and large, a field that takes as its first principle that representations and processes are not quantitative but are comprised of symbols and rules (Chomsky, 1965, 1981, 1995; Marcus, 2001). Over the last 30 years there have been many changes in theories—changes in individual theories (e.g., in Chomskian theories), and the flourishing of many other kinds of theories—but most of them share this one characteristic: that they are not inherently statistical, probabilistic, or quantitative, but rather presume that the medium of representation and the nature of linguistic processes involve symbols and rules.

An extremely interesting exception—perhaps more accurately, an extremely interesting approach to this very issue—is the work of Prince and Smolensky (2004) on Optimality Theory. In this approach, there is still a nonstatistical type of representation—a set of rules and principles that always apply and that apply universally. However, by having a ranking system by which the principles interact with each other, one gets effects that in other theories arise from probabilistic or quantitative interactions among soft nonsymbolic tendencies.

Aside from OT, though, most of Linguistics has taken a nonquantitative approach as a methodological assumption—and in a moment I will say that this has also been a claim about the nature of cognition.

Psycholinguistics has undergone a striking change on this issue over the past 30 years. In the field of Psychology, everything is probabilistic; it is the way one grows up thinking as a psychologist. Even perceiving a light or a tone is not thought of as a discrete event but rather as a probabilistic phenomenon, at the core of which is Signal Detection Theory. In contrast to Linguistics, this probabilistic nature of cognition is not thought of as a performance problem that one could separate from knowledge, which is inherently nonquantitative. Rather, as a psychologist one thinks about a system that is inherently probabilistic in responding to stimulation, with natural variability in input and output as well as in storage mechanisms. These probabilistic characteristics are conceptualized as the real, true underlying nature of the system. Thirty years ago, psycholinguistics stood between these two traditions: one way of thinking about things from the linguistic point of view and the other from the psychological point of view. In the past 30 years, much of the field has been focused on the tension between rules and symbols versus statistics

and quantitative, probabilistic phenomena. See, for example, movement in many parts of psycholinguistics from rules to connectionism to statistical learning (Aslin, Newport & Saffran, 1998; Elman, 2009; Marcus, 2001; Marcus, Vijayan, Bandi Rao, & Vishton, 1999; Mehler, Peña, Nespor, & Bonatti, 2006; Saffran, Aslin, & Newport, 1996; Seidenberg, 1997; Seidenberg, MacDonald, & Saffran, 2003). Today, there is still a tension in the field, indeed opposition, with some people claiming that there are rules and others claiming that there are statistics.

Computational Linguistics has undergone some interesting changes on this same dimension, with symbolic AI dominating previously but much recent work (though not all, of course) being statistical (Charniak, 1993; Eisner, 2002). Some important recent research in computational modeling and computational linguistics takes a Bayesian approach that combines or creates a hybrid of these two types, or an approach using Expectation Maximization that also involves their combination (e.g., comparing symbolic grammars by assessing the probability that linguistic data might be produced by one grammar vs. another) (Eisner, 2002; Goldwater, Griffiths, & Johnson, 2009).

In formal linguistics, the notion that knowledge is made of symbols and rules is not just a methodological approach. Perhaps the most interesting aspect of Chomskian linguistics has been the notion that this is a claim about the nature of the mind: that underlyingly the mind is not probabilistic, that cognition comprises symbols and rules (Chomsky, 1965). The controversy surrounding this claim still divides parts of linguistics from much of psycholinguistics, again with an opposition between the two approaches. There are, however, many ways that differing parts of the field think about statistics versus rules. Some investigators characterize statistics versus rules as different types of computation, while some have argued that there are different cognitive modules or distinct learning mechanisms for the two (a statistical learning mechanism and a rule learning mechanism; cf. Marcus et al., 1999; Pena, Bonatti, Nespor, & Mehler, 2002). Some have proposed hybrids or dual kinds of representations (Pinker, 2000), while others have argued that they are more unified. I have suggested that there may be a sharpening process during learning, one that takes the statistics of sounds and words as the input for learning, but (at least in children) sharpens and regularizes the outcome so that the product behaves more like a rule (Hudson Kam & Newport, 2009; Newport, 1999).

A continuing question for the future is how humans maintain what appear to be these two different types of knowledge. I would suggest that there are some kinds of performance that exhibit one and the other at the same time. People are clearly sensitive to element frequency, bigram frequency, conditional probabilities, and more—not only for language but for most of what they perceive and learn: an amazing array of statistical aspects of the input they experience. At the same time, they also behave in a symbolic way—and (especially children) look like they formulate rule systems, obey principles, and form integrated systems of knowledge (Newport, 1999; Singleton & Newport, 2004; Trueswell & Gleitman, 2007; Wonnacott, Newport, & Tanenhaus, 2008). One of the challenges for the future is to figure out how to integrate these types of knowledge in our descriptions of cognition, rather than argue about them.

3. Modularity

A related issue that I want to mention more briefly, an issue of interest throughout cognition but perhaps nowhere so centrally as in the study of language, is *modularity*. This issue arises in the study of language in two forms: First, is language different from nonlinguistic cognition? That is, is language itself a modularized cognitive function? Second, within language, are there distinct and modularized components of linguistic knowledge and processing? That is, for example, is phonology separate from syntax and semantics? And if modularized, are there fundamentally different kinds of representations and operations that characterize each of those domains? In terms of processing, are there processes that operate on these types of information in strictly sequential fashion, or do they all combine and interact simultaneously?

Thirty years ago, there were widely held notions in the field such as “speech is special,” and most researchers believed that language was different and distinct from other cognitive functions (Lieberman, 1970; Fodor, 1983). It still is the case that people in some parts of the field talk about *the language module*, or *UG* (the acronym referring to a modularized kind of knowledge of language). But a great deal of the field has moved to thinking about interactive constraints on linguistic performance, and about linguistic structure arising from cognitive constraints on learning and real-time processing (Bever, 1970; Hawkins, 1994, 2007; Seidenberg, 1997; Tanenhaus & Trueswell, 2005; Tanenhaus & Brown-Schmidt, 2008). Again, I think these are issues that need to be resolved and brought together in the future.

4. A methodological point

A very brief mention of one change in our field that is methodological: It is surprising to remember how much time we spent 30 years ago in psycholinguistics looking at individual words and printed text. In contemporary psycholinguistics, much of the field now investigates real-time sentence and discourse processing. There are more eye trackers per square foot in my department than one can possibly imagine. Psycholinguists, computational linguists, and formal linguists all do corpus analyses of real speech. A funny example that came to mind as I was preparing this paper is that the basis for what are now called the Brown Corpus and the Penn Treebank was, when I was a graduate student, originally called Kucera and Francis. Kucera and Francis’s (1967) volume was the output of a cadre of graduate students sifting through voluminous amounts of text so that psycholinguists would have word frequency norms for controlling experimental materials. The aim of the project, indeed, was a published word frequency list. In more recent times, this enterprise has been turned on its head. The focus has become the massive texts from which the word frequency counts were derived, rather than the word frequencies themselves; these texts have been digitized and syntactically labeled, and have become the basis of much current-day corpus analysis. That is an interesting shift also.

5. Language and the brain

The last issue I want to focus on, which I think will be much of the future of the language sciences, is: how the brain is organized with respect to language. This has not been the primary focus of the last 30 years in the study of language. Of course there has been some work on language and the brain for many decades, but I think it is fair to say that this has not been a main focus of the field. This is in part because language is the privileged domain of humans, so the most revealing approaches of cellular/molecular or systems neuroscience have not been available for the study of language. But more recently, with fMRI, MEG, NIRS, and other imaging techniques available and widely used, there are methods that are beginning to stimulate many researchers—including many who did not previously work on language and the brain—to start thinking about the problem.

The issue I want to close with is that I think we need to think carefully and in novel ways about what might be the reasonable hypotheses for how language is organized in the brain. This in turn raises a more general question about localization of function for higher cognitive systems. If one looks at any standard neuroscience textbook, one can find depictions of the localization of function for those systems in the brain that are fairly well understood. In the sensory and motor systems, one finds clear organizational patterns for localization of function, with topographic maps that display a patterned layout of the world of stimulation (e.g., the visual field) or the world of motor output (e.g., the hand and arm) onto localized and adjacent pieces of the brain. Even tonotopic auditory cortex, which does not have a spatial mapping to the outer world, is organized in a patterned way, with tone frequencies marching down primary auditory cortex.

How do we develop hypotheses about the neural organization for language, or for any higher cognitive system, if we take these as our best examples of what we know about the brain? What would be candidate hypotheses for language? Many people have thought that the modules of a linguistic grammar would be mapped onto locations of the brain, with a spot for syntax and a spot for semantics (Friederici, 2002). That might be true, but it should not be the only hypothesis that we are thinking about. (Indeed, in my reading of the literature, it is not working out that well so far, with neural activation often quite widespread throughout the left hemisphere language areas for many different types of linguistic tasks.) Perhaps there is a dictionary that runs down the temporal lobe. There actually is quite a bit of interesting evidence that there are spots in the temporal lobe that are involved with tool words as contrasted with animal words (Caramazza & Mahon, 2006). But it does not seem very likely that words from A to Z in the dictionary will be organized alphabetically down the temporal lobe. We ought to be thinking carefully and broadly about what the best hypotheses might be for how language is organized in the brain. David Plaut and Marlene Behrman also spoke at this Cognitive Science Society meeting about some of these issues, suggesting that localization of cognitive functions might arise not from the inherent localization of cognitive modules such as language or face perception, but rather from the interaction of the multiple cognitive or perceptual processes that underlie the task of interest. I want to second their general point: We need some new ways of thinking about how language might be organized in the brain, and

also some consideration of how the layout of other perceptual and cognitive functions might play a role in shaping the topography of language in the brain. We also need new ways of thinking about, and testing, how neural circuitry might accomplish the kinds of generalization and symbolic processes that language entails. While there are some approaches that have taken on this important problem, researchers who focus on the representational side of language have often not been part of the enterprise. My wish for the future is that we might collaborate on addressing these problems of utmost mutual interest.

6. Summary and conclusions

In sum, my agenda for the future would be that we must continue to think about how to integrate rules and statistics rather than to conceptualize them as opposing issues; and we must think in new ways about how the brain might be organized in higher cognitive systems. In addition, we need to address how neural circuits might compute and represent the kind of information relevant to language, concepts, and other aspects of high-level cognition. I hope in the next Cognitive Science Symposium, 30 years hence, we will have solved these simple problems and can stew about some new ones.

Acknowledgments

During the preparation of this paper my work was supported by NIH grant DC00167.

References

- Aslin, R. N., Saffran, J. R., & Newport, E. L. (1998). Computation of conditional probability statistics by 8-month-old infants. *Psychological Science*, *9*, 321–324.
- Bever, T. G. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and development of language* (pp. 279–362). New York: Wiley.
- Caramazza, A., & Mahon, B. Z. (2006). The organization of conceptual knowledge in the brain: The future's past and some future directions. *Cognitive Neuropsychology*, *23*, 13–38.
- Charniak, E. (1993). *Statistical language learning*. Cambridge, MA: The MIT Press.
- Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: The MIT Press.
- Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht, The Netherlands: Foris Publications.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: The MIT Press.
- Eisner, J. (2002). Introduction to the special section on linguistically apt statistical methods. *Cognitive Science*, *26*, 235–237.
- Elman, J. L. (2009). On the meaning of words and dinosaur bones; Lexical knowledge without a lexicon. *Cognitive Science*, *33*, 1–36.
- Fodor, J. (1983). *The modularity of mind*. Cambridge, MA: The MIT Press.
- Friederici, A. D. (2002). Toward a neural basis of auditory sentence processing. *Trends in Cognitive Sciences*, *6*, 78–84.

- Goldwater, S., Griffiths, T. L., & Johnson, M. (2009). A Bayesian framework for word segmentation: Exploring the effects of context. *Cognition*, *112*, 21–54.
- Hawkins, J. A. (1994). *A performance theory of order and constituency*. New York: Cambridge University Press.
- Hawkins, J. A. (2007). Processing typology and why psychologists need to know about it. *New Ideas in Psychology*, *25*, 87–107.
- Hudson Kam, C. L., & Newport, E. L. (2009). Getting it right by getting it wrong: When learners change languages. *Cognitive Psychology*, *59*, 30–66.
- Kucera, H., & Francis, W. N. (1967). *Computational analysis of present-day American English*. Providence, RI: Brown University Press.
- Liberman, A. M. (1970). The grammars of speech and language. *Cognitive Psychology*, *1*, 301–323.
- Marcus, G. F. (2001). *The algebraic mind: Integrating connectionism and cognitive science*. Cambridge, MA: MIT Press.
- Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, *283*, 77–80.
- Mehler, J., Peña, M., Nespors, M., & Bonatti, L. (2006). The soul of language does not use statistics: Reflections on vowels and consonants. *Cortex*, *42*, 846–854.
- Newport, E. L. (1999). Reduced input in the acquisition of signed languages: Contributions to the study of creolization. In M. DeGraff (Ed.), *Language creation and language change: Creolization, diachrony, and development* (pp. 161–178). Cambridge, MA: MIT Press.
- Peña, M., Bonatti, L., Nespors, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, *298*, 604–607.
- Pinker, S. (2000). *Words and rules: The ingredients of language*. New York: Perennial.
- Prince, A., & Smolensky, P. (2004). *Optimality theory: Constraint interaction in generative grammar*. New York: Blackwell.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
- Seidenberg, M. S. (1997). Language acquisition and use: Learning and applying probabilistic constraints. *Science*, *275*, 1599–1604.
- Seidenberg, M. S., MacDonald, M. C., & Saffran, J. R. (2003). Are there limits to statistical learning? *Science*, *300*, 51–52.
- Singleton, J. L., & Newport, E. L. (2004). When learners surpass their models: The acquisition of American Sign Language from inconsistent input. *Cognitive Psychology*, *49*, 370–407.
- Tanenhaus, M. K., & Brown-Schmidt, S. (2008). Language processing in the natural world. In B. C. M. Moore, L. K. Tyler & W. D. Marslen-Wilson (Eds.), *The perception of speech: From sound to meaning. Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*, 1105–1122.
- Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. L. Millers & P. D. Eimas (Eds.), *Handbook of perception and cognition, Vol. 11: Speech, language, and communication* (pp. 217–262). San Diego, CA: Academic Press.
- Trueswell, J. C., & Gleitman, L. R. (2007). Learning to parse and its implications for language acquisition. In G. Gaskell (Ed.), *Oxford handbook of psycholinguistics* (pp. 635–656). Oxford, England: Oxford Univ. Press.
- Wonnacott, E., Newport, E. L., & Tanenhaus, M. K. (2008). Acquiring and processing verb argument structure: Distributional learning in a miniature language. *Cognitive Psychology*, *56*, 165–209.