Language is a social institution

Robert Port
Indiana University
10/28/2009
Draft submitted to *Ecological Psychology*

Abstract
Language seems to pose a problem for ecological psychology. The problem is what kind of things linguistic objects, such as *phonemes* and *words*, might be? Speech, as well as written language, seems to be part of our physical environment, but language also seems to persuasively provide us all with abstract mental structures like words. In a recent special issue of this journal (Hodges and Baron, 2007) it was argued that the social aspects of ecological psychology have not received the attention they deserve since many aspects of motor and perceptual behavior reflect social processes. This essay will reinforce these observations about social perception and social action and further emphasize that societies are complex systems within which the speaker is simply one agent among many (Beckner et al, 2009). Language is essentially social and is probably not definable in terms of any individual psychological system. Language is a part of the culture of human communities that is shaped over historical time. Evidence for the structures of language can be found in the speech corpus but the structures are not physical, invariant tokens used psychologically by individual speakers (as in the traditional view). So it is essential to think about a community and its culture as comprising a *complex system* that is independent of each individual member of the community. This is a critical step toward making sense of language. Many physical (and mathematical) systems have been found in which patterns are *self-organized* (see Haken, et al, 1985; Grossberg 1978). Human communities appear to be such self-organizing systems and they create community-level emergent structures such as the perceptual categories (like *tree, house, wren*, etc.), but also linguistic categories like individual words, idioms and phoneme-like sound patterns. This all means that linguistic objects are essentially social objects or social institutions. They are conventional, statistical patterns in a high-dimensional space created by a speaker community. This system can only be approximately discrete – and is not composed from a restricted set of discrete physical tokens (rather like letters), as was formerly thought. Graphic letters provide a practical model for describing speech, but they must not be confused with the richly detailed patterns that speakers learn, each in their own way.

Linguistic objects or linguistic structures, such as words, idioms, phonemes and phonetic features, seem to pose a problem for ecological psychology. Are they simply physical things or events in the environment? Or should they be viewed as internal, cognitive ``units'' of some mysterious sort? Support for the first hypothesis lies in the fact that physical records of speech sounds can be understood almost as well as live, in-person performances of language. Thus, there must be some kind of effective acoustic form to a language. But it also seems that words and phonemes have a phenomenally vivid, mental or cognitive form, thus suggesting they exist as cognitive objects. This paper will

propose an alternative account for their phenomenal clarity assuming that they too are inherently social features of the environment created by our speech community and do not necessarily have cognitive form. In order to make this argument, it is important to review the standard story about language endorsed by linguists and psychologists in recent centuries and sketch some of the evidence against this view. Then a new approach will be proposed, one that seems quite compatible with ecological psychology and distributed language but which may require some expansion of the notion of ``environment''.

In recent years, the Distributed Language Group has coalesced, endorsing ideas that seem quite compatible with the characterization of language presented here. The term `distributed' seems intended to emphasize that a grammar (or linguistic structure) is not located in each individual, but is distributed across a population (Cowley, this issue). The DLG community (see other papers in this volume) emphasizes that language cannot be separated from the rest of human behavior. Similarly, I will argue that the standard definition of what constitutes ``language'' is strongly biased by our written language experience. In fact, I claim no sharp theoretical distinction can be made between linguistic conventions and other cultural conventions such as gestures, facial expressions, etc. Our professional linguistic definition of `language' is largely constrained to describe the aspects of human vocal interactions that happen to be preserved by an alphabetical orthography. Thus hand and body gestures and facial expressions, etc. are definitely outside linguistics. Intonation seems an awkward borderline case for many linguists. But, in my view, there is no sharp way to distinguish linguistic from nonlinguistic aspects of behavior. The conventions about all those domains are acquired by the same kind of imitation learning: of linguistic expressions, tone of voice, idioms, constructions (Goldberg, 2006) word pronunciations (Labov, 2006), intonation patterns, hand gestures and some facial expressions of the people around us. Because each speaker discovers a way to talk appropriately as they mature, individual speakers will surely find idiosyncratic solutions to the problem of talking like their compatriots. These idiosyncratic grammars are one of the reasons why linguists should focus their attention on language as distributed across a community. One will not generally find linguistic generalizations represented physically within an individual, although they do exist in the corpus. If one wanted to study the detailed representational methods of a single speaker, it might be possible, but then you are not likely to find any of the generalizations that linguists are most interested in.

I will first review the traditional view of speech and language, criticizing it as I go, then sketch the ways that literacy may have shaped our linguistic intuitions.

Traditional View: Language as a mental code

The standard idea about language for at least the past century is that it consists of discrete sound units composed into discrete words which are, in turn, composed into sentences. All mainstream linguistic theories of the 20th C, from Saussure (1916) to Chomsky (1965), proposed variants of these ideas. Linguists differed in how the set of phones or

phonemes was to be determined and what their detailed properties were, but practically all agreed that consonant and vowel tokens are the basic units of language and that they are discrete, nonoverlapping and represent abstract speech sounds. Then words (or morphemes) are the next larger units which, in turn, are combined into utterances whose basic structure consists of phrases and sentences. If such theories have any implications for the psychological form of language, they clearly imply many predictions about the details of speech production and perception. Unfortunately, almost none of these predictions are supported by data (Port and Leary, 2005).

For example, if words are spelled from a discrete alphabet, then (1) speech variation and historical sound changes should exhibit jumps from one sound type to a neighboring sound type, e.g., a [t] might change to a [d] or an [i] to and [ɪ] (or in smaller discrete steps if a larger phonetic alphabet is used). But it has long been known that universal phonetic discreteness is not found and that almost all of the parameters of speech vary continuously, both within individual speakers and between similar dialects. Similarly, (2) there should be directly observable physical correlates of the segmental units (consonants and vowels) that we hear phenomenally listening to a speech audio signal. But it has been known since the first speech waveform displays were made (in early 20th C) that the letter-sized units of speech cannot be distinguished in the speech pressure wave or in spectrographic displays. Generally, it is very difficult even to say, from looking at a sound spectrogram, how many segments there are in some stretch of speech. So the segments we hear phenomenally are not anything like ``segments'' in the sound itself. It should also be true that (3) each phoneme (or segmental speech sound) should exhibit the same acoustic pattern each time it occurs (Chomky and Miller, 1963; Pisoni, 1997). But it has been clear since the 1950's that, for example, stop consonants differing in place of articulation are differentiated from each other by acoustic cues whose form depends on both the preceding and following vowel (Liberman, et al., 1957,1968). In fact, it is generally the case that the letter-like units of speech present different acoustic patterns in nearly every context they occur in. Thus there is nothing in the physical sound to justify or explain using the same consonant in *Dee* as *do*. Another problem is that (4) although phones and phonemes (i.e., letter-like representations of speech) can only support timing relations expressible in terms of the number and serial order of segments, many subtleties of speech timing (most of which speakers have no awareness of) are exploited in languages of the world (Klatt, 1976; Hawkins and Nguyen, 2004; Port and Leary, 2005), studies of speech acoustics and speech perception strongly agree that speakers need detailed speech representations in memory. So, from the viewpoint of what speakers listen to and what they control in speech production, an alphabetical representation would seem to be hopelessly impoverished and inadequate.

Finally, (5) the hypothesis that words are represented in memory using an abstract, speaker-independent alphabet predicts that memories for linguistic material should be speaker-independent and lacking any timing detail. That is, when listening to a list of auditorily presented words, the speech should be remembered in an abstract linguistic code, i.e., using something equivalent to transcription in a phonetic or phonemic alphabet. Of course, speaker-specific information can be stored as well, but speaker idiosyncrasies should not be part of the word representation itself. Speaker properties

could only have an *association* with the linguistic representation of the words. But in recognition memory experiments using lists of spoken words, it has been shown that people recognize that a word has been repeated better when the voice is identical to its first occurrence than if the voice is different (e.g., Goldinger, 1996; Palmeri et al., 1993). One interpretation is that a fairly detailed auditory representation is routinely stored. But also we should be able to make it more difficult to associate the voice with the word by using many different voices instead of just two in the list. If participants tried to remember the voice to improve recognition of the word, then they should be able to do that much better with two voices than with 20. Yet the performance benefit when words are repeated in the same voice is the same no matter how many voices were used. This surprising result (at least for a linguist like me) implies that speakers must normally store auditory speech material in some rich and detailed form that includes speaker-specific information. This conclusion is compatible with the notion of `exemplar' memory or `episodic' memory, as many experiments on vision have found support for (Nosofsky, 1986; Pisoni, 1997; Goldinger, 1996). Such richness of word representations seems to be the generalization one draws across all the evidence cited above. Rich temporal and spectral detail is apparently what is required to do effective speech perception.

These arguments are elaborated elsewhere (Port et al, 1995; Port and Leary 2005: Port 2008, 2009) but, in summary, there little to no experimental support for a cognitive role for abstract, segment-sized units that are invariant across speakers, across speaking rates and across variation in neighboring contexts. However, the likelihood that speech uses a rich and detailed memory code pushes any type of representation based on phones or phonemes right out of the picture. They appear to be irrelevant to speech production and perception. Yet this is very difficult to accept. Linguists and psychologists have resisted drawing this inference for well over a half century now. Why is the inference that there is no real-time psychological role for phones or phonemes so difficult to draw? One reason might be an intuition that speakers would not ignore the availability of an efficient, low bit-rate code like the alphabet when it seems so obvious and readily available. But, of course, it is not ``readily available'' until literacy skills are developed (Ziegler and Goswami, 2005). It may seem transparent to us that *tap* has 3 ``speech sounds'' and the *trap* has 4, however this way of counting speech sounds is the result of learning and practicing reading. But probably the most important reason we have not drawn the inference is that our intuitions about language overwhelmingly testify that a word consists of segmental, letter-like parts, even though it does not either acoustically, articulatorily or in memory or perception.

A written word like *tomato* is usually pronounced by me using the IPA phonetic alphabet as approximately:

$$[t^{h}əme^{I}ɾo^{U}].$$

That is, the initial [t] is aspirated, the second and third vowel are diphthongs (ie, defined in part by their movement as suggested by the superscript vowel symbols) and the second orthographic T is almost invariably pronounced as a tap or flap in my dialect (spelled phonetically as [ɾ] here). Notice that this transcription preserves the same 6 segments as the orthographic word, but it also suggests with superscripts the dynamical gestures speakers are to make when pronouncing it. Although many different pronunciations are

possible for this orthographic word, they can all be approximately described using some configuration of the phonetic alphabet. Both the orthographic representation and the technical phonetic alphabet representation seem highly appropriate and at least approximately correct. But, since the evidence shows speakers do not use such a representation, could there be something that biases us to insist on using letters as the basic components of language? The answer is yes.

Literacy and Linguistics

It may be difficult for us to recall, but every person reading this page spent several hours a week for many years learning to read and to refine their reading skills. It would be naïve to imagine that all this focused mental effort over at least 20 years would have no consequences for our intuitions about the structure of language. Yet we linguists have never paid much attention to the possibility of biased intuitions in our interpretations of speech. It seems likely we overlooked this potential problem due to our lifelong training using alphabet technology in our orthography. Only a few linguists have considered that our orthographic alphabet skills might dominate our perception of speech (see Faber, 1992; Linell, 2005). Chomsky, on the other hand, insisted that we should trust our intuitions completely on matters of grammar and phonology and that doing linguistics is primarily a matter of explicating our intuitions about linguistic structure and finding formal notations for them (Chomsky,1965). It is very easy for us to think about language using an alphabet and it is very difficult to think about the sound of speech in any terms other than alphabetical ones.

But alphabetical writing is a technology which achieved roughly its modern form about 3000 years ago. The earliest Greek alphabet might have been created by a Greek with some Phoenician education who wanted to apply Phoenician writing techniques to Greek. The Phoenician alphabet itself was the culmination of 4-5000 years of earlier record keeping technologies in the middle east that were gradually getting easy enough to learn and convenient enough to use (Fischer, 2005). One major consequence of the development of literacy in the middle east was the growth of the institution of schooling for teaching literacy to children. Schools have been a generally increasing part of life in literate nations ever since (Olson, 1994). Parents in some communities begin teaching literacy to children as young as 2 by playing with alphabet blocks.

Alphabetical writing is certainly very useful, but letters are artifacts. It is difficult to learn to interpret letter sequences as syllables and syllables as letter sequences, so we start teaching children as young as possible. Of course, all of us who have had years of experience reading find the interpretation of continuous speech as a sequence of letter-like units to be very easy and completely natural. One might even ask: how could we expect to hear speech in any other way given all that practice?

Does this mean that phones and phonemes are illusory? Not at all. It just means they are not essential participants in the realtime processing of language, either on input or output. What is illusory here is the idea that phones and phonemes must play a functional role in any cognition involving language because our conscious thoughts about

language use letters. It is probably more accurate to say that phonemes play a functional role for the community. They are regularities or symmetries in the speech corpus of a community which speakers learn to reproduce in their own speech (child stats). How these approximations to a low bitrate code provide benefit to the community and how they benefit individual speakers are two important questions requiring further research (although there are various attempts in print to spell out the benefits of phonology and grammar, Studdert-Kennedy, 2003).

Categories.

One aspect of language as a set of communication conventions is categorization. But it is important not to confuse categories with symbols or formal tokens. *Formal tokens*, like the letters [A-Za-z] and digits [0-9], are discretely different from each other and can be assumed to be recognized and produced with near-perfect accuracy. A *category* is a set of things considered the same by some community – no matter what the reason for calling them the same may be. Presumably, a category name serves as a kind of attractor within semantic space. Cultural conventions chop up the world into parts, many of which have names. A big part of learning a culture and a language is learning the society's ``things'' and all the relevant word labels (Heft, 2007; Hodges and Baron, 2007; Sloutsky, 2003). There are many cultural categories that that shape our behavior that do not have words assigned to them by the popular culture, but may have been explicated by some academic subculture. These include the various categories that seem to constrain linguistic behavior. Academic linguistics have frequently proposed universal status for these grammatical units, but it is far more likely that each culture uses somewhat different grammatical categories (e.g., roughly *noun, verb*, etc.) and phonological categories (e.g., */o/, /n /*, *[±voiced], mora*, etc.) (see. It is important to keep in mind that, as social categories, the detailed specifications will vary from language to language, and, in fine detail, from speaker to speaker (Evans and Levinson, in press). The speech products produced by speakers may fall by convention into various categories, but the speaker usually does not have explicit knowledge of what the categories are. The categories of everyday life, e.g., the names of animals, plants, everyday artifacts, etc, are social conventions and usually depend on language and gesture to demonstrate them. Categories may be defined by a rule (e.g., definition of a *square*) or by any other means, including a simple list of arbitrary members (Smith & Medin, 1981; Glushko et al., 2008). The regularities of a language are in the speech that each child learns from (Tomasello, 1999), and the speaker adapts to the language regularities in some idiosyncratic way depending on their personal history of exposure to languages, dialects and various speakers.

Thus far I have tried to show that the kind of mental representations of language that are actually used by speakers should be thought of as more like richly detailed descriptions of specific heard utterances than like abstract, written form. These concrete events, however, are categorized in rich ways by the culture/language, related to what we call semantics, phonology and syntax. The traditional idea that linguistic patterns could really be simply defined in terms of letter-like physical tokens, discretely ordered in time, was actually a projection of the properties of our traditional alphabet technology onto our

understanding of human psychological processes. The next section will try to present a basic overview of a theory of language as a social institution.

Linguistic Culture: Language as Conventional Regularities of Speech

Homo sapiens seem to have found ways to allow human communities to specialize and to develop complex cultures that create various technologies. Clearly, human sociability contributed to the development of language (Tomasello, 1999) along with the ability to learn statistical regularities after presentation of complex patterns (Newport and Aslin, 2004). Human communities are complex systems capable of adapting their culture over generations (e.g., Beckner, et al, 2009; Smith and Kirby, 2008; Port, 2007; Hruschka et al., 2009; Schoenemann, 1999). These social systems have many agents who interact with each other in complex ways. Human communities are able to evolve rich cultures that are transmitted through the generations (Smith et al., 2003). These systems often organize themselves into various kinds of structures that can be characterized as community organizational structures, religious structures, economic methods, language and many other institutions.

A special trait of humans seems to be our prolific attempts to *categorize* everything around us. One of the most important ``technologies'' created by human communities is spoken language. A language is shaped over generations with respect to the categorization provided by the lexicon and grammar as well as to the range of sound categories the speakers of a community employ (Hock and Joseph, 1996). But there will always be uncertainty about what the patterns actually are, due to unavoidable uncertainty about such issues as ``who is a member of this community and who is not." It is not even clear how to distinguish words from ``common phrases," from idioms, from ``ways of talking about things". Our orthography makes arbitrary decisions about what counts as a *word* (and is separated from others by a space) vs. a *phrase* of *sentence* but each speaker must find their own way to store memories of language and generate new utterances. In the next section, I will sketch some of the main components of linguistic description.

Linguistic Structures

What are some of the structures of language that seem to be created by communities of speakers? There is, first, a ``lexicon,'' a list of word-like chunks (even though a precise list is impossible as well as the precise specification for each word) and, second, so-called ``grammatical regularities,'' illustrated by case markings on nouns and tense markings on verbs. Also languages always seem to employ a restricted inventory of so-called ``speech sounds.'' The problem is that our use of letters biases us to interpret ``speech sounds'' as letter-sized units, like a consonant or vowel. But actually speech sound is not countable in any consistent way. It provides a continuous dynamic gesture and cannot be divided into segment-sized pieces (Browman and Goldstein, 1992). The importance of the notion of ``restricted inventory'' is that words tend to resemble each other quite a bit, such that, for most speech fragments, other fragments can be found that are almost identical.

Each language explores a very limited range of the possible human speech sounds. Also, sometimes pairs of words differ from each other in very similar ways. For example, the distinction between the 6 words in Set 1 below from each other is fairly similar to the differences between the 6 words of Set 2. The orthographic notations here are intended to stand for continuous speech gestures or their continuous-time acoustic description. As noted above, some acoustic differences will be the same between *ban-pan* and *Bill-pill*, although the formant transitions from the stops into the vowel will be different. But the third set, where the stops occur in a different position in the syllable, presents more of a problem. The voicing distinction between [b] and [p], etc. is manifested quite differently between *Bill-pill* and *lab-lap* since the vowels in the words ending in [b,d,g] are quite a bit longer than the vowels in the words ending in [p,t,k] and the aspiration that distinguished *Bill* from *pill* occurs nowhere.

ban pan	Bill pill	lab lap
Dan tan	dill till	lad *lat
*gan can	gill kill	lag lack
Set 1	Set 2	Set 3

But English orthography implicitly claims the same consonantal ``speech sounds'' are used in all 3 sets of words, even though research on the physical speech signals shows that speakers of the language must store and represent the detailed auditory form of the words. Presumably the language has a good reason for favoring word specifications involving only partial similarity between syllable-initial and syllable-final contexts.

This is an example of symmetrical patterns created by a community of speakers. These words show symmetry since, for example, the ``voicing property'' is treated as the same for all ``sounds'' in the same column in the sets above and the place of articulation feature is identical for all stops in the same row. Of course, this is true only sometimes since, for example, (a) the contrast between /t/ and /d/is largely neutralized before an unstressed vowel, as in *budding/butting* in my dialect, and (b) the difference between /b,d,g/ and /p,t,k/ does not exist after an /s/ (cf. *spot, stow, Scot*, but not *sbot, sdot, sgot*) although the claim of identity across contexts may be approximately true in many cases. The very simple and discrete regularities suggested by the alphabetical notation in the table above apply only to letters or other graphic tokens on paper and not to any speech sound. Actual phonologies are only approximately discrete and require far more degrees of freedom than letters do.

Literacy.

What has happened is that recently, long after developing language perhaps as much as a half million years ago, certain human communities developed a practical and easily teachable alphabetic writing system. The approximately low-dimensional statistical patterns in human speech shown above can be mapped reasonably well onto a short list of alphabet tokens (Olson, 1994; Harris, 2000). Thus texts could be created. Over the past 3k years, there has been some progress in making teaching more effective, but the basic methods for teaching literacy have hardly changed: memorize the alphabet, then start learning short common words first. The alphabet literacy we all share trains us to phenomenally experience speech as comprised of a sequence of letters from a very small

alphabet, and to consist of discrete *words* and *sentences* as well. The fit between the alphabet and the real phonology of a language is good enough that we literates have been happy to overlook the problems (since we learned to just `get over' the inconsistencies and the arbitrariness while still very young and learning to read). In the past century, formal mathematical grammars have taken the properties of orthographic letters (discreteness, serial order, etc.) and generalized them to create mathematical systems (Harrison, 1978) that led to computers and to theories about how language could really be formal represented in the human brain (Chomsky and Miller, 1967; Chomsky, 1965).

The model for thinking about language that I am proposing is that individual speakers must learn the patterns of speech of their community in a form that affords both understanding the speech of others and also affords appropriate speech production. Each speaker stores some memory representations of heard speech that include both fairly rich sensory data and various categorizations of the speech along with any generalizations they have been able to draw (e.g., about word identities, phoneme identities, and semantic categories). Each speaker does this independently based upon somewhat different sensory experience, so the cognitive representations will have to differ.

The linguistic community provides exposure to fragments of a linguistic culture with a long history (somewhat different fragments in each case, of course). The linguistic culture of a community provides a massive set of conventional categorizations of the experience of that community. These include both unnamed categories and ones that have a lexical name or a description that requires only a few component names. The speaker learns ``speech chunks'' by rote example at first (Grossberg, 1978; 2003) but gradually comes to categorize speech the way their family and neighbors do (Werker and Tees, 1984) and learns to control their own body to produce speech that conforms to the conventions of their community.

Conclusions.

It seems to me that the radical story I am endorsing here is compatible with most of the data dealing with language and its psychological basis. It also seems to be compatible with the ideas of the distributed language group since I emphasize the vast differences between written and spoken language as well as the inseparability of linguistic from nonlinguistic social behavior. Since we are all literate, we fail to see that our thoughts about language (whether we are lay or professional) have been powerfully shaped by our reading and writing experience. Although almost all linguists have followed Saussure (1916) in claiming to be studying the spoken language not the written language, the fact is that we, like Saussure, never really escaped from letter-based characterizations of language. When we think of `words', `speech sounds' and `sentences' in our descriptions of language, we are importing the conventions of our writing system and trying to use them technically. This is the `language stance' that Cowley warns of (Cowley, this issue). Confusion about the distinction between the written language and the spoken language creates difficulties for many types of research (see Love's, 2004, criticisms of Clark, 2001).

What I have tried to show here is that the neat code-like units of our normal conscious thoughts about language – that is, the letter-like speech sounds, the discrete words, morphemes, prepositional phrases, full sentences, etc. – play at most a very tiny, specialized role in real-time conversational performance. These linguistic patterns only exist at the level of the community, as a kind of social institution. Linguistics should stop trying to describe the representations of abstract linguistic units stored in people's heads. These patterns do not exist in the form of abstract, formal memories stored somewhere. Instead, our every utterance is a creative behavioral response to our experience, shaped by every possible facet of our lives, but generally compatible with the conventions of the language. A language is, first of all, the set of linguistic conventions about speech shared by some community. For several millennia some communities have also had a system of writing conventions based partly on the phonetic value of letters (Harris, 2000). These too should also be viewed as part of a language, but they differ in fundamental ways from what are used for speech.

References

Beckner, C., Blythe, R., Bybee, J., Christiansen, M., Croft, W., Ellis, N., et al. (2009). Language is a complex adaptive system: Position paper. *Language Learning, 59*(1), 1-27.

Browman, C., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica, 49*, 155-180.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, Massachusetts: MIT Press.

Chomsky, N., & Miller, G. (1963). Introduction to the formal analysis of natural languages. In R. D. Luce, R. R. Bush & E. Galanter (Eds.), *Handbook of Mathematical Psychology* (Vol. 2, pp. 323-418). New York: Wiley and Sons.

Clark, A. (2001). *Mindware: An introduction to the philosophy of cognitive science*. New York: Oxford University Press.

Cowley, S. (in press, 2000). The language stance. *Ecological Psychology*.

Evans, N., & Levinson, S. (2009 to appear). The myth of language universals: Language diversity and its importance for cognitive science. *Brain and Behavioral Sciences*.

Faber, A. (1992). Phonemic segmentation as epiphenomenon: Evidence from the history of alphabetic writing. In P. Downing, S. Lima & M. Noonan (Eds.), *The Linguistics of Literacy* (pp. 111-134). Amsterdam: John Benjamins.

Fischer, S. (2005). *A history of writing*: Reaktion Books.

Glushko, R., Maglio, P., Matlock, T., & Barselou, L. (2008). Categorization in the wild. *Trends in Cognitive Science, 12*(4), 129-135.

Goldberg, A. (2006). *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.

Goldinger, S. D. (1996). Words and voices: Episodic traces in spoken word identification and recognition memory. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*, 1166-1183.
Grossberg, S. (1978). A theory of human memory: Self-organization and performance of sensory-motor codes, maps and plans. In R. Rosen & F. Snell (Eds.), *Progress in Theoretical Biology* (Vol. 5, pp. 233-374). New York: Academic Press.
Grossberg, S. (2003). The resonant dynamics of speech perception. *Journal of Phonetics, 31*, 423-445.
Haken, H., Kelso, J. A. S., & BUnz, H. (1985). A theoretical model of phase transitions in human hand movements. *Biological Cybernetics, 51*, 347-356.
Harris, R. (2000). *Rethinking Writing*. London: Continuum.
Harrison, M. (1978). *Introduction to Formal Language Theory*. Reading, Mass.: Addison-Wesley.
Hawkins, S., & Nguyen, N. (2004). Influence of syllable-final voicing on the acoustic onset of syllable-onset /l/ in English. *Journal of Phonetics, 32*, 199-231.
Heft, H. (2007). The social constitution of perceiver-environment reciprocity. *Ecological Psychology, 19*(2), 85-105.
Hock, H. H., & Joseph, B. (1996). *Language History, Language Change and Language Relationship* (2d ed.): MNouton de Gruyter.
Hodges, B., & Baron, R. (2007). On making social psychology more ecological and ecological psychology more social. *Ecological psychology, 19*(2), 79-84.
Holland, J. (1995). *Hidden Order: How Adaptation Builds Complexity*. Cambridge, Mass: Perseus Books.
Hruschka, D., Christiansen, M., Blythe, R., Croft, W., Heggarty, P., Mufwene, S., et al. (2009). Building social cognitive models of language change. *Trends in Cognitive Science, 13*(11), 464-469.
Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America, 59*, 1208-1221.
Labov, W., Ash, S., & Boberg, C. (2006). *The Atlas of North American English*. Berlin: Mouton-de Gruyter.
Liberman, A. M., Delattre, P., Gerstman, L., & Cooper, F. (1968). Perception of the speech code. *Psychological Review, 74*, 431-461.
Liberman, A. M., Harris, K. S., Hoffman, H., & Griffith, B. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology, 54*, 358-368.
Linell, P. (2005). *The Written Language Bias in Linguistics*. Oxford: Routledge.
Love, N. (2004). Cognition and the language myth. *Language Sciences, 26*, 525-544.
Newport, E. L., & Aslin, R. (2004). Learning at a distance I: statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*, 127-162.
Nosofsky, R. (1986). Attention, similarity and the identification-categorization relationship. *Journal of Experimental Psychology: General, 115*, 39-57.
Olson, D. R. (1994). *The World on Paper: The Conceptual and cognitive implications of writing and reading*. Cambridge: Cambridge University Press.
Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology, Learning, Memory and Cognition, 19*, 309-328.

Pisoni, D. B. (1997). Some thoughts on `normalization' in speech perception. In K. Johnson & J. Mullennix (Eds.), *Talker variability in speech processing* (pp. 9-32). San Diego: Academic Press.

Port, R. (2007). How words are stored in memory: Beyond phones and phonemes. *New Ideas in Psychology, 25*, 143-170.

Port, R. (2009). The dynamics of language. In R. Meyers (Ed.), *Encyclopedia of Complexity and System Science* (pp. 2310-2323). Heidelberg: Springer-Verlag.

Port, R., Cummins, F., & McAuley, D. (1995). Naive time, temporal patterns and human audition. In R. Port & T. v. Gelder (Eds.), *Mind as Motion: Explorations in the Dynamics of Cognition* (pp. 339-371). Cambridge, Mass.: Bradford Books/MIT Press.

Port, R. F., & Leary, A. (2005). Against formal phonology. *Language, 81*, 927-964.

Saussure, F. d. (1916). *Course in General Linguistics* (W. Baskin, Trans.). New York: Philosophical Library.

Schoenemann, P. T. (1999). Syntax as an emergent characteristic of the evolution of complexity. *Minds and Machines, 9*, 309-346.

Sloutsky, V. (2003). The role of similarity in the development of categorization. *Trends in Cognitive Sciences, 7*(6), 246-251.

Smith, E., & Medin, D. (1981). *Categories and Concepts*. Cambridge, Mass: Harvard University Press.

Smith, K., Brighton, H., & Kirby, S. (2003). Complex systems in language evolution: The cultural emergence of compositional structure. *Advances in Complex Systems, 6*(4), 537-558.

Smith, K., & Kirby, S. (2008). Cultural evolution: implications for understanding the human language faculty and its evolution. *Philosophical transactions of the royal society B, 336*, 3591-3603.

Studdert-Kennedy, M. (2003). Launching language: The gestural origin of discrete infinity. In M. H. Christiansen & S. Kirby (Eds.), *Language Evolution: The State of the Art*. Oxford, UK: Oxford University Press.

Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, Mass: Harvard University Press.

Twaddell, W. F. (1935). On defining the phoneme. *Language, Monograph*(16).

van Gelder, T., & Port, R. (1995). Its about time. In R. Port & T. v. Gelder (Eds.), *Mind as Motion: Explorations in the Dynamics of Cognition* (pp. 1-44.). Cambridge, Mass.: MIT Press.

Werker, J., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development, 7*, 49-63.

Ziegler, J., & Goswami, U. (2005). Reading acquisition, developmental dyslexia and skilled reading across languages: A psycholinguistic grain size theory. *Psychological Bulletin, 131*, 3-29.