

## **Language Has Evolved to Depend on Multiple-Cue Integration**

*Morten H. Christiansen*\*

Department of Psychology, Cornell University, Ithaca, NY, USA

Santa Fe Institute, NM, USA

R. Botha & M. Everaert (Eds.) (2013). The evolutionary emergence of language: Evidence and inference (pp. 42-61). Oxford: Oxford University Press.

**Word count** (main text only): 6,114

\*Please address correspondence to:  
Morten H. Christiansen  
Department of Psychology  
228 Uris Hall  
Cornell University  
Ithaca, NY 14853  
Phone: 607-255-3834  
e-mail: [christiansen@cornell.edu](mailto:christiansen@cornell.edu)

## Abstract

Recent research on language evolution suggests that language to a large extent has been shaped by cultural evolution to fit domain-general constraints deriving from the human brain. An implication of this view is that much of the neural hardware involved in language is not specific to it. This means that language has to be acquired largely by mechanisms that are not uniquely dedicated for this purpose. But how could language have evolved to be as complex as it is without language-specific constraints? Here, I propose that language has evolved to rely on a multitude of probabilistic information sources for its acquisition, allowing language to be as expressive as possible while still being learnable by domain-general learning mechanisms. In particular, I focus on the probabilistic contributions of phonological information to the learning of basic aspects of syntax. Based on results from corpus analyses, computational modeling and human experimentation, I argue that a probabilistic relationship exists between what a word sounds like and how it is used: nouns tend to sound like other nouns and verbs like other verbs. Importantly, these sources of phonological information, or cues, not only play an important role in language acquisition but also affect syntactic processing in adulthood. I conclude that the integration of phonological cues with other types of information is integral to the computational architecture of our language capacity. Such multiple-cue integration, in turn, is what has allowed language to evolve to be learnable without language-specific constraints given the rich sources of information available in the input.

## Introduction

A growing body of work suggests that language may have evolved primarily by way of cultural evolution rather than biological adaptation. Evidence in support of this perspective on language evolution comes from computational modeling (e.g., Chater, Reali & Christiansen, 2009; Kirby, Dowan & Griffiths, 2007), behavioral experimentation (e.g., Kirby, Cornish & Smith, 2008; Fay, Garrod & Roberts, 2008), linguistic analyses (e.g., Evans & Levinson, 2009; Heine & Kuteva, 2007), and many other lines of scientific inquiry (see Chater & Christiansen, in press, for a review). A key hypothesis emerging from this work is that language has been shaped by cultural evolution to fit domain-general constraints deriving from the human brain (Christiansen, 1994; Deacon, 1997—see Christiansen & Chater, 2008, for a review of the historical pedigree of this perspective). An implication of this view is that much of the neural hardware involved in language is not specific to it (though see Christiansen, Reali & Chater, 2011). That is, language has to be acquired largely by mechanisms that are not uniquely dedicated for this purpose. But how could language have evolved to be as complex as it is without language-specific constraints?

In this chapter, I propose that language has evolved to rely on a multitude of probabilistic information sources for its acquisition, allowing language to be as expressive as possible while still being learnable by domain-general mechanisms (see also, Christiansen & Chater, 2008; Chater & Christiansen, 2010). Crucially, though, this perspective does not deny the existence of genetic constraints on language but instead questions the presupposition that these necessarily have to be linguistic in nature. This proposal has important theoretical ramifications for how we should think about language, as discussed in the first section, *Language as a Culturally Evolved Linguistic System*. Next, in *The Sound of Syntax*, I introduce the notion of multiple-cue

integration in language acquisition and processing, using the role of phonological information in syntax learning as an extended case study. Finally, implications are drawn from this perspective for how we should construe the species-specificity of language and cross-linguistic similarities in *Biological Adaptation and Language Universals*.

### **Language as a Culturally Evolved Linguistic System**

A key question for language evolution research is to explain why language is the way it is, and how it got to be that way. The cultural evolution perspective suggests that the structure of language derives primarily from processes of cultural transmission involving repeated cycles of learning and use, constrained by the properties of the human brain. Thus, instead of asking, “Why is the brain so well-suited for learning language?”, we need to turn the question upside-down and ask, “Why is language so well-suited to being learned by the brain?” This shift in focus allows us to see that the fit between the mechanisms employed for language and the way in which language is acquired and used may be explained by processes of cultural evolution shaped by the human brain. Thus, similar to the proposed cultural recycling of prior cortical maps for recent human innovations such as reading and arithmetic (Dehaene & Cohen, 2007), I suggest that language likewise has evolved by ‘piggy-backing’ on preexisting neural substrates, inheriting their structural constraints. These constraints—including socio-pragmatic considerations, the nature of our thought processes, perceptuo-motor factors, as well as cognitive limitations on learning, memory and processing—subsequently have been amplified and embedded in language through cultural evolution (Christiansen & Chater, 2008). In this way, cultural transmission—both vertically (across generations) and horizontally (within generations)—has shaped language to be as learnable and processable as possible by largely

domain-general mechanisms. Indeed, this process of cultural evolution has similarities with biological evolution, as already noted by Darwin (1871) in *The Descent of Man*: “The formation of different languages and of distinct species, and the proofs that both have been developed through a gradual process, are curiously the same.” (p. 59). Further underscoring the parallel between processes of biological and linguistic change, Darwin concludes that “The survival and preservation of certain favored words in the struggle for existence is natural selection.” (p. 60-61).

This emerging view of language evolution suggests that language is best construed as a culturally evolved linguistic system. Consequently, what has evolved is *not* a set of neural structures specific to language; rather, cultural evolution produces *a system of linguistic constructions* specific to a given speech community (i.e., a language). In formal linguistic terms, these constructions can be viewed within a construction grammar framework as conventionalized form-meaning mappings (e.g., Croft, 2001; Goldberg, 2006). In psychological terms, the constructions can be interpreted from a usage-based perspective as the basic linguistic building blocks of language learning and processing (e.g., Tomasello, 2003; Christiansen & MacDonald, 2009). Importantly, though, these constructions do not exist in a vacuum but have been shaped by cultural evolution to be easy to learn and use in the context of the other constructions that make up the linguistic system of a given speech community (see Christiansen & Chater, 2008, for further discussion). Language, as culturally evolved linguistic system, comes to reflect the domain-general constraints of its learners/users and can therefore not be fully understood without taking this process of linguistic adaptation into account.

An intriguing theoretical implication of this viewpoint is that it promises to simplify the problem of language acquisition: The learner’s biases will be the right biases because language

has been optimized by past generations of learners to fit those very biases (Chater & Christiansen, 2010; Zuidema, 2003). However, this does not entail that language acquisition is trivial. Children still face the formidable task of learning the multitude of constructions that make up the language of their speech community. Moreover, children must also discover how these constructions relate to one another. Fortunately, the child does *not* approach this task as a *tabula rasa* learner. Instead, children can rely on the very same domain-general constraints, ranging from socio-pragmatic expectations (e.g., Levinson, 2000) to cognitive/memory limitations (e.g., O’Grady, 2005), to which language has been shaped by previous generations of language learners/users. Over developmental time, children then develop neural representations relevant for using the linguistic system(s) to which they are exposed. Thus, domain-general mechanisms can give rise to language-specific representations, but the latter can only be fully understood when investigated as the outcome of a developmental process shaped by constraints that are not specific to language (Karmiloff-Smith, 2009).

It is important to note that the conception of language presented here is incommensurable with the distinction between internalized language (I-language) and externalized language (E-language), where the latter “construct is being understood independently of the properties of the mind/brain” (Chomsky, 1986: p. 20). Indeed, from the viewpoint of cultural evolution it makes little theoretical sense to distinguish between E-language and I-language because neural linguistic representations cannot be understood in the absence of what they are representations of, and vice versa. On the one hand, the information that children come to represent in their brains about their native language will to a large extent depend on the linguistic input that they receive and the largely domain-general constraints they brought to bear on the acquisition process. Yet, on the other hand, the linguistic input that children get, in turn, has been shaped by

previous learners and users of that particular language, who have similar neural constraints. Thus, the neural representations of language cannot be understood independently of the input that gave rise to them, and likewise, the linguistic system of a speech community cannot be understood independently of the neural constraints that shaped it.

Another theoretical consequence of viewing language as a culturally evolved linguistic system is that language evolution is not conceptually different from language change. Indeed, on this account, the distinction between evolution and change is an artificial one because it presupposes a perspective in which language *evolution* involves the emergence of a biological basis for language, whereas language *change* characterizes subsequent linguistic variation (constrained by the previously evolved language-specific neural substrate). In contrast, the current perspective suggests that it is possible to observe the signature of language evolution in processes of linguistic change. For example, the same grammaticalization processes observed in historical language change are also hypothesized to have played a role in the cultural evolution of language. Thus, the study of grammaticalization can provide insight into what earlier and simpler forms of language may have looked like (Heine & Kuteva, 2007).

The current perspective, however, leaves open the potentially unanswerable questions of language origin: how, when, and why language emerged in the human lineage. Instead, this theory presupposes that at some point in human prehistory language *did* emerge, and that language in its earliest form was quite simple and without much discernable linguistic structure. The aim of the current account, then, is to provide explanations—couched primarily in terms of cultural evolution—of the subsequent emergence of the complex linguistic systems we can observe today. But how could language have become as complex as it is and still be acquired without language-specific constraints? Next, I argue that language has solved this problem by

“recruiting” multiple, partially overlapping sources of information to facilitate its acquisition and use.

### **The Sound of Syntax**

Central to the thesis of language as a culturally evolved linguistic system is the suggestion that this system has been shaped by multiple constraints deriving from the human brain. But might this process of multiple-constraint satisfaction have ramifications beyond cultural evolution, affecting how we acquire and use language? Here, I explore the hypothesis that pre-existing neural constraints not only provided important restrictions on cultural evolution but also made available multiple sources of information—or *cues*—that can facilitate both the acquisition and use of language. By recruiting such cues, some of which may be partially overlapping and redundant, language could evolve culturally to become more expressive, while still being learnable and processable by mechanism not dedicated to language. Consequently, as a product of cultural evolution, every language today incorporates its own unique constellation of probabilistic cues to signal different aspects of linguistic structure from word boundaries (e.g., Mattys, White & Melhorn, 2005) to syntactic relations (Monaghan & Christiansen, 2008) and beyond (Evans & Levinson, 2009). Because these cues are probabilistic in nature, and therefore unreliable when considered in isolation, multiple-cue integration has become a necessary component of language acquisition.

The hypothesis put forward here is thus that cultural evolution has shaped languages to depend on multiple-cue integration for their acquisition and processing. But how might this work? In this section, I put forward an extended case study in multiple-cue integration, focusing

on one possible source of information about syntactic structure—phonological cues—and how these may be integrated with distributional information during language acquisition and processing. First, I present results from evolutionary simulations indicating that cultural evolution can lead to the recruitment of cues to facilitate acquisition. Quantitative results from corpus analyses then demonstrate that phonological and distributional cues can be integrated to provide reliable information about lexical categories. Computational modeling confirms that domain-general learners can successfully utilize such cues. Finally, human experiments establish that children use phonological cues during word learning and that adults cannot help but pay attention to the sound of syntax during sentence processing.

### **Recruitment of Cues during Cultural Evolution of Language**

Christiansen and Dale (2004) conducted a set of simulations to investigate whether cultural evolution could result in the recruitment of cues to facilitate the learning of more complex linguistic structure. As learners, they employed simple recurrent networks (SRNs, Elman, 1990); a type of connectionist model that implements a domain-general learner with sensitivity to complex sequential structure in the input. This model is trained to predict the next element in a sequence and thus learns in a self-supervised manner from violations of its own expectations regarding what should come next. The SRN model has been successfully applied to the modeling of both language processing (e.g., Elman, 1993)—including multiple-integration in speech segmentation (Christiansen, Allen & Seidenberg, 1998) and syntax acquisition (Christiansen, Dale & Reali, 2010)—as well as sequential learning (e.g., Botvinick & Plaut, 2004). As a model of human performance, the SRN has been shown to closely mimic the processing of different kinds of recursive linguistic constructions (Christiansen & Chater, 1999; Christiansen &

MacDonald, 2009) as well as the sequential learning of nonadjacent dependencies (Misyak, Christiansen & Tomblin, 2010). In addition, the SRN has also been applied to the modeling of potential co-evolution between language and learners (Batali, 1994).

The languages, on which Christiansen and Dale's models were trained, were generated by small context-free grammars, each derived from the grammar skeleton illustrated in Table 1. The curly brackets indicate that the order of the constituents on the right-hand side of a rule can be either as is (head-first) or in the reverse order (head-final). The SRNs were expected to use the distributional information afforded by the order of words in the sentences as a cue to the underlying structure of the language. As additional cues to linguistic structure, the languages could recruit a constituent cue and a lexical cue. The constituent cue was an additional input unit that could mark phrase boundaries by being activated following the constituents from a particular phrase structure rule (e.g., N (PP) #, where “#” indicates the activation of the constituent cue after the NP and optional PP). The lexical cue was another input unit that could be co-activated with any of the 24 words in the vocabulary. Thus, there were three potential sources of information for learning about the structure of a language in the form of distributional, constituent, and lexical cues.

Cultural evolution was simulated by having 5 different languages compete against one another, with fitness determined by how easy it was for the SRNs to learn a language. At the beginning of a simulation, 5 different languages were randomly generated based on the grammar skeleton with a random combination of constituent and lexical cues. Each language was then learned by 5 different SRNs, with a language's fitness being computed as the average across the 5 networks. The most easily learned language, along with 4 variations of it, would then form the basis for the next generation of languages, each being learned by 5 networks. Again, the most

easily learned language would be selected as the parent for the next generation, and the process repeated until the same language won for 50 consecutive generations. Language variation was implemented by randomly changing two of the three cues: 1) changing the head-order of a rule, 2) adding or deleting the constituent unit for a rule, or 3) adding or deleting the co-activation of the lexical unit for a word. Ten different simulations were run, each with different initial randomizations.

Of the ten simulations, one never settled but the results of the remaining nine simulations followed a similar pattern. First, all languages ended up with a highly regular head-ordering, with at least 5 of the 6 phrase structure rules being either all head-initial or all head-final. This fits the general tendency for word order patterns in natural languages to be either head-initial or head-final (e.g., Dryer, 1992)<sup>1</sup>. Second, the constituent cue always separated NPs from other phrases, consistent with evidence from corpus analyses indicating that prosodic cues, such as pauses and pitch changes, are used to delineate phrases in both English and Japanese child-directed speech (Fisher & Tokura, 1996). Finally, the lexical cue reliably separated word classes, with six of the runs resulting in the lexical cue separating function words from content words. This is similar to the acoustic differentiation of function and content words observed in English (Cutler, 1993). To place these results in context, it is important to note that given the combination of the three different cues in these simulations, there were nearly three-quarters of a million<sup>2</sup> different possible linguistic systems that could have evolved through cultural evolution. Thus, it is not a trivial result that these simulations culminated in linguistic systems that incorporate properties closely resembling those of natural language.

The simulations by Christiansen and Dale (2004) suggest that linguistic systems can recruit cues to facilitate learning when undergoing cultural evolution. The integration of these

cues, in turn, allows language to become more complex while still being learnable by domain-general mechanisms. If natural language is a culturally evolved multiple-cue integration system, as indicated by these simulations, it should be possible to uncover such cues in current languages and to quantify their usefulness. Next, I report on corpus analyses focusing on two of the cues explored in the evolutionary simulations: lexical information in the form of the phonology of individual words and distributional information in the form of word co-occurrence patterns.

### **Quantifying the Usefulness of Phonological and Distributional Cues**

Do the phonological forms of words contain information relevant for syntax acquisition? The standard assumption of the arbitrariness of the sign (de Saussure, 1916) might be taken to suggest otherwise. Indeed, it has been argued that it is a universal characteristic of human language that the relationship between the form of a word and its meaning is arbitrary (Hockett, 1960). This assumption is fundamental to most modern grammatical theories on both sides of the Chomskyan divide. For example, Pinker (1999, p. 2) states that “onomatopoeia and sound symbolism certainly exist, but they are asterisks to the far more important principle of the arbitrary sign—or else we would understand the words in every foreign language instinctively, and never need a dictionary for our own!” In a similar vein, Goldberg (2006, p. 217) notes that “... the particular phonological forms that a language chooses to convey particular concepts [...] generally are truly arbitrary, except in relative rare cases of phonaesthemes.” However, the simulations by Christiansen and Dale (2004) suggest that the sign may not be entirely arbitrary; rather, a systematic relationship should exist between the sound of a word and its lexical category, if the phonological form of a word is to be useful for syntax acquisition as a lexical cue.

In a series of corpus analyses of child-directed speech, Monaghan, Chater and Christiansen (2005) quantified the potential usefulness of phonological cues to lexical categories. More than five million words were extracted from the CHILDES database (MacWhinney, 2000), comprising more than a million utterances spoken in the presence of children. Phonological forms and lexical categories were gleaned from the CELEX database (Baayen, Pipenbrock & Gulikers, 1995) and results reported for the 5,000 most frequent words. As potential cues to lexical categories, Monaghan et al. used 16 different phonological properties (listed in Table 2) that have been proposed to be useful for separating nouns from verbs (and function words from content words). Instead of treating each cue in isolation, the 16 cues were combined into a unified phonological representation for each word. A statistical analysis<sup>3</sup> was then conducted using these representations, resulting in decent classifications of both nouns (58.5%) and verbs (68.3%)—with an indication that phonological cues may be more useful for discovering verbs than nouns. The advantage of phonological cues for verbs was subsequently confirmed by further analyses in Christiansen and Monaghan (2006).

Importantly, though, because the phonological cues are probabilistic in nature they did not predict the lexical category of a word perfectly. As an additional cue, Monaghan et al. therefore assessed the usefulness of distributional information using a simple, developmentally plausible approach. They selected the 20 most frequent words in the corpus (*are, no, there, this, your, that's, on, in, oh, do, is, and, I, that, what, to, a, it, the, you*) and recorded how often these preceded one of the target words (e.g., *you want*). The rationale was that even though the child may not know the meaning of the 20 context words, these word forms nonetheless constitute highly frequent acoustic events to which the child is likely to be sensitive. To determine the usefulness of the distributional patterns thus recorded, Monaghan et al. used an information-

based measure<sup>4</sup> to assess the strength of the association between the context word and the target word. The distributional information relevant to a specific target word was then represented as a unified representation containing the associative strength scores for each of the 20 context words (e.g., for *cats*, the scores for *are cats*, *no cats*, *there cats*, and so on). These distributional cue representations were then submitted to the same statistical analysis as the phonological cues. The results showed a very good classification of nouns (93.7%) but not of verbs (31.1%).

The results from the two analyses suggest that the usefulness of phonological and distributional cues may differ across nouns and verbs. Perhaps integration across the two types of cues may improve classification? Monaghan et al. combined the phonological cues and the distributional cues into a combined representation<sup>5</sup>, redid their analyses, and obtained reliable classifications of both nouns (67.0%) and verbs (71.4%). When considering correct classifications of nouns and verbs together, they further noted an interesting interaction of phonological/distributional cues with frequency, as shown in Figure 1a. Distributional cues appear to work very well for high-frequency words but less so for low-frequency words. This is likely because high-frequency words occur in more contexts and this provides for more accurate distributional information about their lexical categories. Phonological cues, on the other hand, seem to work better for low-frequency words than for high-frequency words. This may be explained by the tendency for high-frequency words to be shortened, perhaps leading to the omission of important phonological cues to their lexical category. In contrast, low-frequency words are not subjected to the same shortening pressures, allowing the cues to remain in place. Crucially, though, when the two types of cues are integrated, good classification can be found across all frequency bins, as illustrated in Figure 1b.

The results presented so far apply only to English. If language, in general, has evolved to rely on multiple-cue integration, then it should be possible to find evidence of similar kinds of cue information available in other languages as well. However, many of phonological cues used by Monaghan et al. (2005) were specific to English (Table 1) and thus may not work for other languages. Monaghan, Christiansen and Chater (2007) therefore generated a set of 53 cross-linguistic phonological cues, including gross-level word cues such as length, consonant cues relating to manner and place of articulation of phonemes in different parts of the words, and vowel cues relating to tongue height and position as well as whether the vowel was reduced. They then conducted analyses of child-directed speech in English, French, and Japanese. Using the new cues, they replicated the results of the previous study in terms of correct noun/verb classification (16 cues: 63.4% vs. 53 cues: 67.5). Noun/verb classification using phonological cues was also very good for both French (82%) and Japanese (82%). Classification performance was further improved across all three languages (English: 94%; French: 91.4%; Japanese: 93.4%) when the phonological cues were integrated with distributional cues (computed as before).

Together, the results of the corpus analyses show that across representatives of three different language genera—Germanic (English), Romance (French), and Japanese—child-directed speech contains useful cues for distinguishing between nouns and verbs (see also, Kelly, 1992). The results are thus consistent with the hypothesis that, as a result of the cultural evolution of language, words contain within them, the sound of syntax: nouns and verbs differ in terms of their phonology<sup>6</sup>. Importantly, the specific cues differed considerably across languages, suggesting that each language has recruited its own unique set of cues to facilitate acquisition through multiple-cue integration. However, these analyses only demonstrate that there are

probabilistic cues available for learning about aspects of syntax. Next, we shall see that a domain-general sequential learner, the SRN, can take advantage of both phonological and distributional cues to learn about syntax.

### **Multiple-Cue Integration by a Sequential Learner**

A potential concern regarding multiple-cue integration is that there are many kinds of information that could potentially inform language acquisition. As noted by Pinker (1984, p. 49), “... in most distributional learning procedures there are vast numbers of properties that a learner could record, and since the child is looking for correlations among these properties, he or she faces a combinatorial explosion of possibilities. ... Adding semantic and inflectional information to the space of possibilities only makes the explosion more explosive.” Pinker expresses a common intuition about the use of multiple, partially correlated sources of information by a domain-general learning device: that combining different kinds of partially reliable information can only result in unreliable outcomes. However, research in formal learning theory has shown that this intuition is incorrect. Mathematical analyses of neural network learning using the Vapnik-Chervonenkis Dimension<sup>7</sup> have shown that multiple-cue integration with correlated information sources will not lead to a combinatorial explosion but instead to improved learning (Abu-Mostafa, 1993). This holds even when one or more of the cues are either uncorrelated or otherwise uninformative with respect to the acquisition task, in which case they have no negative effect on learning (see Allen & Christiansen, 1996, for neural network applications, including to the SRN). Thus, mathematically speaking, Pinker’s intuitive concern about combinatorial explosion is unfounded.

Although the issue of combinatorial explosion is not a problem in principle, it may nonetheless pose a considerable obstacle in practical terms. Christiansen and Dale (2001) sought to address this issue head-on by training SRNs to do multiple-cue integration, given a corpus of artificially generated child-directed speech. The corpus incorporated declarative, imperative, and interrogative sentences with subject-noun/verb agreement and variations in verb argument structure. In one simulation, the networks were provided with three partially reliable cues to syntactic structure (word length, lexical stress, and pitch change) and three cues not related to syntax (presence of word-initial vowels, word-final voicing, and relative speaker pitch). The results of the simulations indicated that the SRNs were able to ignore the three unrelated cues while taking full advantage of informative ones, as indicated by the mathematical analyses (see Gogate & Hollich, 2010, for a discussion of how language learners may detect invariances in the input more generally).

The question remains, though, whether Christiansen and Dale's SRN model can scale up to deal with the kind of cues found in the corpus analyses described previously. To answer this question, Reali, Christiansen and Monaghan (2003) trained SRNs on a full-blown corpus of natural speech directed at children between the ages of 1 year and 1 month to 1 year and 9 months (Bernstein-Ratner, 1984). Each word in the input was encoded in terms of the 16 phonological cues used in the Monaghan et al. (2005) corpus analyses (and shown in Table 2). Given a word represented in terms of these phonological cues, the task of the networks was to predict the next lexical category in the utterance. Thus, the network would receive both phonological cues (in terms of the 16 cue-representations for each word) and distributional cues (in terms of the co-occurrence of words in the corpus). To assess the usefulness of the phonological cues relative to the distributional ones, a second group of networks was also

trained. For these networks, the phonological-cue representation for a given word was randomly reassigned to a different word to break the correlations between phonology and lexical category. The results showed that the SRNs trained to integrate phonological and distributional cues performed significantly better than the networks provided only with distributional cues. Further analyses of the networks' internal states indicated that the phonological cues were particularly useful for processing novel words, allowing the network to place itself in a "noun state" when processing novel nouns and in a "verb state" when encountering new verbs.

The results of the SRN simulations suggest that a domain-general sequential learner can learn aspects of syntactic structure via multiple-cue integration. Despite intuitions to the contrary, a combinatorial explosion does not occur. Rather, the right cues are recruited to facilitate acquisition because the language has evolved to be learnable by way of those very cues. For example, phonological cues promote better learning and better generalization to new words. To be able to take advantage of these cues, children become attuned to the relevant cues in their native language during the first years of life, as we shall see next.

### **Phonological Cues in Acquisition and Processing**

The corpus analyses indicated that there are useful phonological cues for language acquisition and the SRN simulations demonstrated that a sequential learner can take advantage of them—but are children sensitive to phonological cues when exposed to new words? Storkel (2001, 2003) have shown that preschoolers find it easier to learn novel words when these consist of phonotactically common sound sequences. However, these studies did not address the question of whether children may use phonological cues to learn about the syntactic role of words. Fitneva, Christiansen and Monaghan (2009) therefore conducted a word learning study to investigate whether children implicitly use phonological information when guessing about the

referents of novel words. To create novel words that were either noun-like or verb-like in their phonology, Fitneva et al. used a measure of phonological typicality, originally proposed by Farmer, Christiansen and Monaghan (2006). Phonological typicality measures how typical a word's phonology is relative to other words in its lexical category, and reliably reflects the phonological coherences of nouns and verbs (Monaghan, Christiansen, Farmer & Fitneva, 2010). Thus, noun-like nouns are typical in terms of their phonology of the category of nouns, and likewise verb-like verbs are phonologically typical of other verbs. When asking English monolingual second-graders to guess whether a novel word referred to a picture of an object or a picture of an action, the children used the phonological typicality of the nonword in making their choices. Interestingly, as predicted by the corpus analyses (Christiansen & Monaghan, 2006), verbs benefitted more from phonological cues than nouns.

It may be objected that second-graders—as language learners—are too “old” to serve as a suitable population with which to investigate the usefulness of phonological cues, especially if such cues are to be used to inform early syntactic acquisition. To address this objection, Fitneva et al. conducted a second study with another group of second-graders, who were enrolled in a French immersion program. The stimuli were the same as in the experiment with the monolingual children. Crucially, though, whereas half the nonwords were verb-like and the other half noun-like with respect to English phonology, all the nonwords were noun-like according to French phonology. Two groups of the French-English bilingual children were tested, with the only difference being in the language used for the instructions. When given English instructions, the bilingual children behaved exactly like the monolingual English children, showing an effect of English phonological typicality. However, when the instructions were provided in French, the patterns of results changed, in line with French phonology. Hence, not only did the children seem

to use phonological cues to make guesses about whether a novel word is a noun or a verb but they were also able to do so after relatively little short experience with the relevant phonology (less than two years of exposure for the children in the French immersion program).

The results of the word learning study suggest that phonological cues may come into play early in syntax acquisition. Farmer et al. (2006) explored whether multiple-cue integration involving phonological cues extend into adulthood. Using the measure of phonological typicality, they demonstrated that the processing of words presented in isolation is affected by how typical their phonology is relative to their lexical category: noun-like nouns are read aloud faster, as are verb-like verbs. Similarly, Monaghan et al. (2010) showed that people are faster to make lexical decisions about whether a presented item is a real word or not, if that word is phonologically typical of its lexical category. Farmer et al. further showed that the phonological typicality of a word could even affect how easy it is to process in a sentence context. Indeed, for noun/verb homonyms (e.g., *hunts* as in *the bear hunts were terrible...* versus *the bear hunts for food...*), if the continuation of the sentence is incongruent with the phonological typicality of the homonym, then people both experience on-line processing difficulties and have problems understanding the meaning of the sentence.

Together, the results of the human experimental studies indicate that the use of phonological cues during acquisition is so important that it becomes a crucial part of the developing language processing system. The phonological properties of words facilitate lexical acquisition through multiple-integration and become an intricate part of lexical representations. As consequence, adult language users cannot help but pay attention to phonological cues to syntactic structure when processing language.

### **Biological Adaptation and Language Universals**

In this chapter, I have proposed that language has evolved by way of cultural evolution to exploit multiple cues so as to be maximally expressive while still being learnable by domain-general mechanisms. Evolutionary simulations were discussed, indicating how language may recruit cues to facilitate learning. A prediction from this perspective on the cultural evolution of language is that each language should have its own constellation of cues. Cross-linguistic corpus analyses have confirmed this prediction with regard to phonological cues, showing that the relationship between a word's sound and how it's used is not arbitrary. Computational simulations have demonstrated that domain-general sequential learners can take advantage of phonological cues in the context of multiple-cue integration—as can children when learning new words. The adult sentence processing results further indicate that the use of phonological cues becomes a crucial part of the emerging language processing system. Of course, phonological cues are not the only useful sources of information for learning about aspects of syntax; rather, phonological cues are integrated with other sources of information during language acquisition, including distributional (e.g., Redington, Chater & Finch, 1998), prosodic (e.g., Fisher & Tokura, 1996), semantic (e.g., Bowerman, 1973), and pragmatic (e.g., Tomasello, 2003) cues (see Morgan & Demuth, 1996; Monaghan & Christiansen, 2008, for reviews). Hence, language has evolved to rely on multiple-cue integration in both acquisition and processing, making it integral to the computational architecture of our language system.

Although the perspective on language evolution advocated here has focused on cultural evolution, it does not preclude that there may have been biological adaptations for aspects of language that would improve language learning and use. For example, simulations investigating the circumstances under which the spectacular diversity of human languages (Evans & Levinson,

2009) can be reconciled with the presumed biological uniformity of their speakers point to biological adaptations for flexible learning to better accommodate rapid cultural evolution of language (Baronchelli, Chater, Pastor-Satorras & Christiansen, 2011). Additional simulation results indicate that there may also have been adaptations for some functional aspects of language (Christiansen et al., 2011), such as the ability to learn a large inventory of form-meaning mappings or better processing of complex sequences. Moreover, it is possible that humans may have evolved more sophisticated abilities for multiple-cue integration than what we might observe in other primates—though little data is currently available to address this question. Thus, on the present account, the reason why extant non-human primates do not have language may be due to humans having gone through several biological adaptations, most of which are not specific to language, but which provided the right kind of perceptuo-motor, cognitive, conceptual, and socio-pragmatic foundations for language to “take off” by way of cultural evolution.

The emphasis on the cultural evolution of language suggests that there may be few true language universals, in the sense of being a linguistic property that holds across all languages (Christiansen & Chater, 2008). If such universals exist, they are likely to derive from the domain-general constraints that have shaped the cultural evolution of language, rather than from a genetically encoded language-specific endowment. Multiple-cue integration, as discussed here, may be a plausible candidate for a universal property common to all languages. Crucially, though, multiple-cue integration is not unique to language but also plays a key role in, for example, vision (e.g., Tsutsui, Taira & Sakata, 2005) and sensori-motor control (e.g., Green & Angelaki, 2010). Another possible universal relating to phonological cues may be a division of labor between arbitrariness and systematicity in word form-meaning mappings, deriving from

opposing pressures from the task of learning the meanings of individual words, on the one hand, and the process of discovering how to use these words syntactically, on the other. Monaghan, Christiansen and Fitneva (2011) present results from computational simulations, human experiments, and corpus analyses indicating that whereas one part of a word's phonological form may have a primarily arbitrary form-meaning mapping to facilitate meaning individuation, another part of the same word tends to incorporate systematicity to assist in the acquisition of lexical category information. Yet other potential universals may derive from the other types of domain-general constraints that shape the cultural evolution of language; e.g., the universality of turn taking in conversations likely deriving from language's role in social interaction (Stivers et al., 2009). More generally, however, the picture that emerges is one in which linguistic diversity becomes a key explanandum for theories of language evolution—as also advocated by Evans & Levinson (2009)—and for which the explanans, I submit, must include the idea that language has evolved to depend on multiple-cue integration.

### Footnotes

<sup>1</sup> Recently, Dunn, Greenhill, Levinson & Gray (2011) have shown that word order correlations tend to be lineage specific rather than universal as proposed by Dryer (1992). The current approach is consistent with these results under the assumption that the history of a particular language provides additional constraints on the specific path along which a language changes.

<sup>2</sup> The number of possible linguistic systems was calculated as follows: there were 6 rules with 2 head-orderings, each with or without the constituent cue, and with each language having a 24-word vocabulary, in which each word could be associated with the lexical cue or not:  $6^4 \times 24^2 = 746,496$ .

<sup>3</sup> Each word was represented by a 16-place vector. This means that each word corresponds to a point in a 16-dimensional space defined by the 16 phonological cues. Monaghan et al. employed a discriminant analysis to determine whether the nouns and verbs formed separate clusters in this phonological space. Informally, this type of statistical analysis inserts a hyper-plane into the 16-dimensional phonological cue space to produce the most optimal separation of nouns and verbs into two different categories. Correct classification of nouns and verbs can then be computed given how well the hyperplane separates the two categories from one another.

<sup>4</sup> Monaghan et al. used a modified version of the Dunning (1993) log-likelihood score to estimate the informational value of the distributional cues. Informally, this measure provides an estimation of how surprising it is that the context and target words occur together given how often each occurs on their own. Each word was then represented as a 20-place vector, corresponding to the signed Dunning log-likelihood scores for each of the 20 context words.

Classification of nouns and verbs given these distributional representations was then assessed using a discriminant analysis.

<sup>5</sup> This simply involved combining the 16-place phonological cue vector with the 20-place distributional vector, resulting in a 36-place multiple-cue vector representation for each word. A discriminant analysis was conducted on the 36-dimensional cue space defined by these word representations.

<sup>6</sup> That the phonological forms of words carry information about their syntactic use as nouns or verbs does not necessarily require the postulation of universal lexical categories. Instead, phonological and distributional cues provide probabilistic information about how words can be used in sentential contexts and this is what is assessed by the corpus analyses reported in this chapter.

<sup>7</sup> The Vapnik-Chervonenkis (VC) dimension establishes an upper bound for the number of examples needed by a learning process that starts with a set of hypotheses about the task solution. A hint may lead to a reduction in the VC dimension by weeding out bad hypotheses and reduce the number of examples needed to learn the solution.

## References

- Abu-Mostafa, Y.S. (1993). Hints and the VC Dimension. *Neural Computation*, 5, 278-288.
- Allen, J. & Christiansen, M.H. (1996). Integrating multiple cues in word segmentation: A connectionist model using hints. In *Proceedings of the 18th Annual Cognitive Science Society Conference* (pp. 370-375). Mahwah, NJ: Lawrence Erlbaum.
- Baayen, R. H., Pipenbrock, R. & Gulikers, L. (1995). *The CELEX lexical database* (CD-ROM). Linguistic Data Consortium, University of Pennsylvania, Philadelphia, PA.
- Baronchelli, A., Chater, N., Pastor-Satorras, R. & Christiansen, M.H. (2011). *The biological origin of linguistic diversity*. Manuscript in preparation.
- Batali, J. (1994). Innate biases and critical periods: Combining evolution and learning in the acquisition of syntax. In R. Brooks & P. Maes (Eds.), *Artificial Life IV*. Cambridge, MA: MIT Press.
- Bernstein-Ratner, N. (1984). Patterns of vowel modification in motherese. *Journal of Child Language*, 11, 557-578.
- Botvinick, M. & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action. *Psychological Review*, 111, 395-429.
- Bowerman, M. (1973). Structural relationships in children's utterances: Semantic or syntactic? In T. Moore (Ed.), *Cognitive development and the acquisition of language* (pp. 197-213). New York: Academic Press.
- Chater, N & Christiansen, M.H. (in press). A solution to the logical problem of language evolution: Language as an adaptation to the human brain. In M. Tallerman & K. Gibson (Eds.), *Oxford Handbook of Language Evolution*. Oxford, UK: Oxford University Press.

- Chater, N. & Christiansen, M.H. (2010). Language acquisition meets language evolution. *Cognitive Science*, 34, 1131-1157.
- Chater, N., Reali, F. & Christiansen, M.H. (2009). Restrictions on biological adaptation in language evolution. *Proceedings of the National Academy of Sciences*, 106, 1015-1020.
- Chomsky, N. (1986). *Knowledge of language*. New York: Praeger.
- Christiansen, M.H. (1994). *Infinite languages, finite minds: Connectionism, learning and linguistic structure*. Unpublished doctoral dissertation, Centre for Cognitive Science, University of Edinburgh.
- Christiansen, M.H., Allen, J. & Seidenberg, M.S. (1998). Learning to segment speech using multiple cues: A connectionist model. *Language and Cognitive Processes*, 13, 221-268.
- Christiansen, M.H. & Chater, N. (1999). Toward a connectionist model of recursion in human linguistic performance. *Cognitive Science*, 23, 157-205.
- Christiansen, M.H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral & Brain Sciences*, 31, 487-558.
- Christiansen, M.H. & Dale, R. (2001). Integrating distributional, prosodic and phonological information in a connectionist model of language acquisition. In *Proceedings of the 23rd Annual Conference of the Cognitive Science Society* (pp. 220-225). Mahwah, NJ: Lawrence Erlbaum.
- Christiansen, M.H. & Dale, R. (2004). The role of learning and development in the evolution of language. A connectionist perspective. In D. Kimbrough Oller & U. Griebel (Eds.), *Evolution of communication systems: A comparative approach. The Vienna Series in Theoretical Biology* (pp. 90-109). Cambridge, MA: MIT Press.

- Christiansen, M.H., Dale, R. & Reali, F. (2010). Connectionist explorations of multiple-cue integration in syntax acquisition. In S.P. Johnson (Ed.), *Neoconstructivism: The new science of cognitive development* (pp. 87-108). New York: Oxford University Press.
- Christiansen, M.H. & MacDonald, M.C. (2009). A usage-based approach to recursion in sentence processing. *Language Learning*, 59 (Suppl. 1), 126-161.
- Christiansen, M.H. & Monaghan, P. (2006). Discovering verbs through multiple-cue integration. In K. Hirsh-Pasek & R.M. Golinkoff (Eds.), *Action meets words: How children learn verbs* (pp. 88-107). New York: Oxford University Press.
- Christiansen, M.H., Reali, F. & Chater, N. (2011). Biological adaptations for functional features of language in the face of cultural evolution. *Human Biology*, 83, 247-259.
- Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. New York: Oxford University Press.
- Cutler, A. (1993). Phonological cues to open- and closed-class words in the processing of spoken sentences. *Journal of Psycholinguistic Research*, 22, 109-131.
- Darwin, C. (1871). *The descent of man, and selection in relation to sex, Vol. 1*. London: John Murray.
- Deacon, T.W. (1997). *The symbolic species: The coevolution of language and the brain*. New York: WW Norton.
- Dehaene, S. & Cohen, L. (2007). Cultural recycling of cortical maps. *Neuron*, 56, 384-398.
- de Saussure, F. (1916). *Course in general linguistics*. New York: McGraw-Hill.
- Dryer, M. (1992). The Greenbergian word order correlations”, *Language*, 68, 81-138.
- Dunn, M., Greenhill, S.J., Levinson, S.C. & Gray, R.D. (2011). Evolved structure of language shows lineage-specific trends in word-order universals. *Nature*, 473, 79-82.

- Dunning, T. (1993). Accurate methods for the statistics of surprise and coincidence. *Computational Linguistics*, 19, 61-74.
- Elman, J. L. (1990). Finding structure in time. *Cognitive Science*, 14, 179-211.
- Elman, J.L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48, 71-99.
- Evans, N. & Levinson, S. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, 32, 429-492.
- Farmer, T.A., Christiansen, M.H. & Monaghan, P. (2006). Phonological typicality influences on-line sentence comprehension. *Proceedings of the National Academy of Sciences*, 103, 12203-12208.
- Fay, N., Garrod, S. & Roberts, L. (2008). The fitness and functionality of culturally evolved communication systems. *Philosophical Transactions of the Royal Society of London (B)*, 363, 3553-3561.
- Fisher, C. & Tokura, H. (1996). Acoustic cues to grammatical structure in infant-directed speech: Cross-linguistic evidence. *Child Development*, 67, 3192-3218.
- Fitneva, S.A., Christiansen, M.H. & Monaghan, P. (2009). From sound to syntax: Phonological constraints on children's lexical categorization of new words. *Journal of Child Language*, 36, 967-997.
- Goldberg, A. (2006). *Constructions at work: The nature of generalization in language*. New York: Oxford University Press.
- Gogate, L.J. & Hollich, G. (2010). Invariance detection within an interactive system: A perceptual gateway to language development. *Psychological Review*, 117, 496-516.

- Green, A.M. & Angelaki, D.E. (2010). Multisensory integration: Resolving sensory ambiguities to build novel representations. *Current Opinion in Neurobiology*, 20, 353-360.
- Heine, B. & Kuteva, T. (2007). *The genesis of grammar: A reconstruction*. New York: Oxford University Press.
- Hockett, C. F. (1960). The origin of speech. *Scientific American*, 203, 89-96.
- Karmiloff-Smith, A. (2009). Nativism versus neuroconstructivism: Rethinking the study of developmental disorders. *Developmental Psychology*, 45, 56-63.
- Kelly, M.H. (1992). Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review*, 99, 349-364.
- Kirby, S., Cornish, H., and Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, 105, 10681-10685.
- Kirby, S., Dowman, M. & Griffiths, T. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences*, 104, 5241-5245.
- Levinson, S. C. (2000). *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (Third Edition). Mahwah, NJ: Lawrence Erlbaum Associates.
- Mattys, S.L., White, L. & Melhorn, J.F. (2005). Integration of multiple speech segmentation cues: A hierarchical framework. *Journal of Experimental Psychology: General*, 134, 477-500.
- Misyak, J.B., Christiansen, M.H. & Tomblin, J.B. (2010). Sequential expectations: The role of prediction-based learning in language. *Topics in Cognitive Science*, 2, 138-153.

- Monaghan, P., Chater, N. & Christiansen, M.H. (2005). The differential role of phonological and distributional cues in grammatical categorisation. *Cognition*, 96, 143-182.
- Monaghan, P. & Christiansen, M.H. (2008). Integration of multiple probabilistic cues in syntax acquisition. In H. Behrens (Ed.), *Trends in corpus research: Finding structure in data* (TILAR Series) (pp. 139-163). Amsterdam: John Benjamins.
- Monaghan, P., Christiansen, M.H. & Chater, N. (2007). The Phonological-Distributional Coherence Hypothesis: Cross-linguistic evidence in language acquisition. *Cognitive Psychology*, 55, 259-305.
- Monaghan, P., Christiansen, M.H., Farmer, T.A. & Fitneva, S.A. (2010). Measures of phonological typicality: Robust coherence and psychological validity. *The Mental Lexicon*, 5, 281-299.
- Monaghan, P., Christiansen, M.H. & Fitneva, S.A. (2011). The arbitrariness of the sign: Learning advantages from the structure of the vocabulary. *Journal of Experimental Psychology: General*, 140, 325–347.
- Morgan, J.L. & Demuth, K. (Eds.). (1996). *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*. Mahwah, NJ: Lawrence Erlbaum.
- O'Grady, W. (2005). *Syntactic carpentry: An emergentist approach to syntax*. Mahwah, NJ: Erlbaum.
- Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
- Pinker, S. (1999). *Words and rules*. New York: Basic Books.
- Realo, F., Christiansen, M.H. & Monaghan, P. (2003). Phonological and distributional cues in syntax acquisition: Scaling up the connectionist approach to multiple-cue integration. In

*Proceedings of the 25th Annual Conference of the Cognitive Science Society* (pp. 970-975).

Mahwah, NJ: Lawrence Erlbaum.

Redington, M., Chater, N. & Finch, S. (1998). Distributional information: A powerful cue for acquiring syntactic categories. *Cognitive Science*, 22, 425-469.

Stivers, T., Enfield, N.J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., de Ruiter, J.P., Yoon, K.-Y. & Levinson, S.C. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, 106, 10587-10592.

Storkel, H.L. (2001). Learning new words: Phonotactic probability in language development. *Journal of Speech, Language, and Hearing Research*, 44, 1321-1337.

Storkel, H.L. (2003). Learning new words II: Phonotactic probability in verb learning. *Journal of Speech, Language, and Hearing Research*, 46, 1312-1323.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

Tsutsui, K., Taira, M. & Sakata, H. (2005). Neural mechanisms of three-dimensional vision. *Neuroscience Research*, 51, 221-229.

Zuidema, W. (2003). How the poverty of the stimulus solves the poverty of the stimulus. In S. Becker, S. Thrun & K. Obermayer (Eds.), *Advances in neural information processing systems 15* (pp. 51-58). Cambridge, MA: MIT Press.

### **Acknowledgments**

Many thanks to Christina Behme and four anonymous reviewers for helpful comments on a previous version of this chapter.

**Table 1: The Grammar Skeleton used by Christiansen & Dale (2004)**

---

S	→	{NP VP}
NP	→	{N (PP)}
NP	→	{N PossP}
VP	→	{V (NP)}
PP	→	{adp NP}
PossP	→	{poss NP}

---

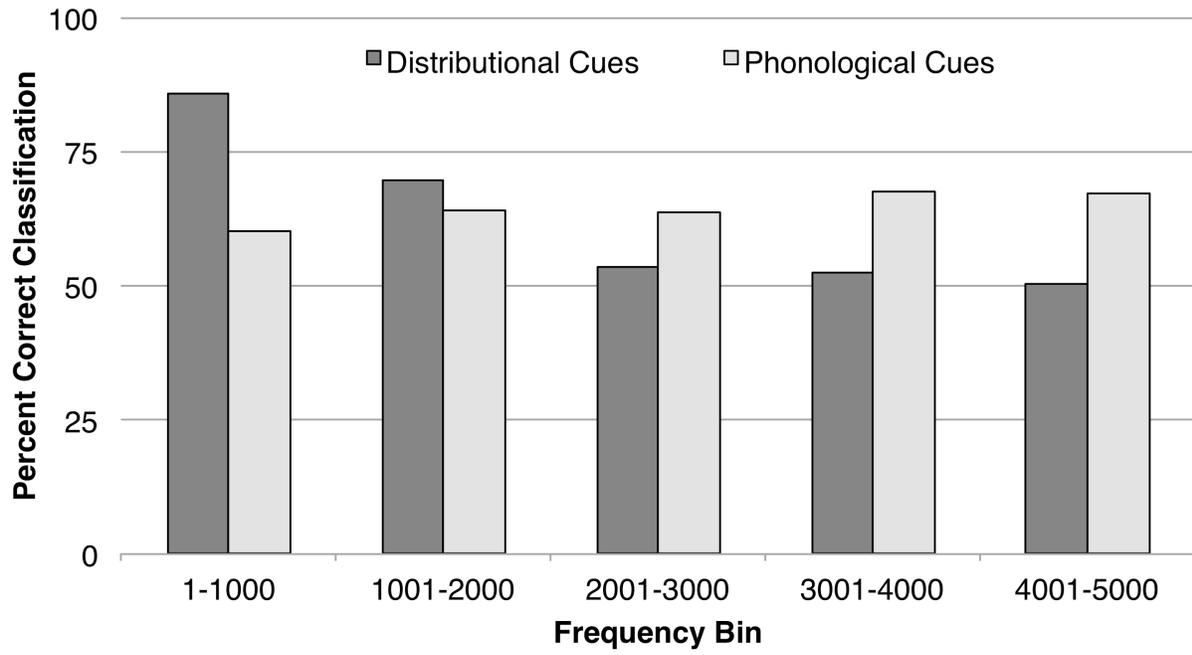
**Table 2: The 16 Phonological Cues used by Monaghan, Chater & Christiansen (2005)**

Phonological Cue	Example: <i>fingers</i>
<b>Word level</b>	
Length in phonemes	6
Length in syllables	2
Presence of stress	1
Syllable position of stress	1
<b>Syllable level</b>	
Number of consonants in word onset	1
Proportion of phonemes that are consonants	0.66
Proportion of syllables containing reduced vowel	0.5
Reduced 1 <sup>st</sup> vowel	0
<i>-ed</i> inflection	0
<b>Phoneme level</b>	
Proportion of consonants that are coronal	0.25
Initial /ð/	0
Final voicing	1
Proportion of consonants that are nasals	0.25
Position of stressed vowel	1
Position of vowels	1.5
Height of vowels	1

### Figure Captions

*Figure 1.* The percentage of nouns and verbs correctly classified as such across different frequency bins for a) distributional and phonological cues treated separately, and b) when both cues are integrated with one another.

a)



b)

