The role of adaptation in understanding linguistic diversity

Gary Lupyan (lupyan@wisc.edu)
Department of Psychology, University of Wisconsin-Madison

Rick Dale (rdale@ucmerced.edu)
Cognitive and Information Sciences, University of California, Merced

1. Introduction

Human populations display a dazzling diversity of cultural practices. Clothing styles, building techniques, cooking practices, art, and legal systems all show enormous variability. Attempts to understand why people in equatorial Africa wear different clothes from people in the Arctic would not go far without considering differences in climate. It is rather obvious that traditional cooking techniques are strongly constrained by the kinds of foods available in the area and the preparation those foods require. Yet, when it comes to language—another culturally transmitted system showing enormous cross-cultural diversity—the assumption most linguists and psychologists have made is that the variability is not meaningfully related to factors that strongly constrain, or even determine, other aspects of human culture.[1] We argue that—just as looking to the physical environment is necessary to explain differences in cultural practices such as clothing styles and building techniques—looking to the social and physical environment is necessary for understanding at least some reasons why languages vary in the way they do.

We begin by addressing a fundamental question of why there is linguistic diversity at all, and suggest that languages diversify in part because they are adaptations to different human environments. Next, we describe our previous work showing that it is possible to account for some specific aspects of linguistic diversity by considering the socio-demographic "niches" in which languages are used. On this view, languages adapt over time to optimize learnability and information-transmission within specific niches. This perspective is largely in line with that proposed by the contributors to this volume, particularly the work of Trudgill, and in the analyses of Martowicz, Burridge, Palmer, Stebbins, and Tadmor.

2. Why are there so many languages?

According to the story of the Tower of Babel, there was a time when all humans spoke a single language. A hubristic attempt to build a tower to the heavens led to God jumbling human languages ("Babel" comes from the Hebrew balal, to jumble). It is instructive to ask why there should be such a story at all. While existential wondering such as "where does the world come from?" (God made it) or "why do all humans look similar?" (made in God's image) seem like

[1] We do not mean to suggest that cultures vary without limit, or that certain environments always produce particular cultural artifacts or institutions. The associations are always probabilistic. For example, it is very unlikely that a culture without access to clay develops tradition of pottery, or that new types of sailing technology are invented in a landlocked environment.

natural material for religious texts to address, the question of why are there so many languages appears rather more esoteric by comparison. One answer is that for most of human history languages were extraordinarily regional (a similar point is made by Trudgill in this volume). From one settlement, one would need to walk a long way to encounter humans with obvious physical (racial/ethnic) differences. In comparison, in most places in the world the distance to the nearest language would have been quite short. Even at present, half of the world's languages have fewer than 7,000 speakers, and half are spoken over an area smaller then Luxembourg (Ethnologue, Gordon, 2005). As a result, people would frequently be exposed to individuals who looked very much like them, and yet spoke different languages, leading them to wonder "why?" Strikingly, we still do not have a clear answer. Explanations of linguistic diversity, both at the dialect and language level, have focused on drift. For example, Sapir (1921) writes:

> …dialects arise not because of the mere fact of individual variation, but because two or more groups of individuals have become sufficiently disconnected to drift apart, or independently, instead of together. So long as they keep strictly together, no amount of individual variation would lead to the formation of dialects. In practice, of course, no language can be spread over a vast territory or even over a considerable area without showing dialectic variations, for it is impossible to keep a large population from segregating itself into local groups, the language of each of which tends to drift independently. (p. 161)

Linguistic drift arising from both synchronic and diachronic processes is undoubtedly important in accounting for typological patterns. But it may not be the only driver of linguistic diversity. Consider an analogous argument that drift is the source of biological variation. We can easily apply Sapir's analysis to, for example, a colony of finches. As the initial finch group splinters, the members of each subgroup will be more likely to mate with one another and, over time, the two groups will drift further apart genetically, eventually producing different species. But such an account is missing a critical element: adaptation. An account of biological diversity that excludes adaptation cannot explain why, compared to the ancestral species, some finch species should come to have wider beaks, while others, longer beaks. Thus, the divergence of the groups is due not just to assortative mating, but also to the groups being subjected to (however slightly) different selective pressures. Even populations that remain in close proximity can rapidly diverge if their members come to occupy distinct niches that place an adaptive pressure on some trait, e.g., beak shape or foraging strategy. Indeed, the Galapagos finches initially studied by Darwin occupied small and often overlapping territories (see Weiner, 1995 for a book-length account of the fascinating research—past and present—on Darwin's finches).

Thus, our ability to explain why a particular animal has the features it does clearly requires a consideration of the environment in which (and in a sense, *for* which) it has evolved. Here, we take this argument into the domain of languages. Just as a beak of some shape can be viewed as an adaptation to a particular environment, so a particular grammar can be, on this account, viewed as an adaptation.

3. How different are languages, really?

Before attempting to answer the question of what environments shape languages and how they might do so, it is worth considering the more fundamental question of whether languages really are different from each other in interesting ways. One cannot dispute that languages differ at least at the surface level, but in some quarters, it has been fashionable to assume that such variability is illusory and that even its study detracts from the "real" goal of understanding the deep structure of language. Such a deep characterization of language is often taken to be the generative model on which all languages are based (i.e., Universal Grammar). For example, Pinker (1994) writes:

> According to Chomsky, a visiting Martian scientist would surely conclude that aside from their mutually unintelligible vocabularies, Earthlings speak a single language. (p. 232).

It is true that all languages share certain design principles such as compositionality and symbolic reference that make them, as a group, distinct from other forms of communication (both non-human animal communication and nonverbal human communication). Insofar as there are universal design features that separate human language from other communication systems, studying these features (e.g., symbolic reference, compositionality) (Deacon, 1997; Hockett, 1966) involves delving into the question of origins—questions that the Chomskyan research program has avoided, for the most part.[2]

Claiming that "Earthlings speak a single language" is a bit like saying that there is only one kind of bird; that apart from different colors and sizes, and shapes, and so on, all birds are the same. It may indeed be useful to distinguish between animals that are birds and those that are not and we can fruitfully ask what is true of *all* birds. But surely it is at least equally sensible to ask why some birds eat fish and others eat insects and what characteristics make a bird suitable for one type of diet versus another, as well as why some parts of the world have many different species of birds and others have few. If we examine languages at the level of analysis that would be of interest to, for example, a comparative biologist studying species of birds, how substantial would we find the differences between languages to be?

---

[2] Consider an observation analogous to Chomsky's Martian scientist: "all life on earth is just variation on a Universal Grammar of DNA; thus, differences among species are just dialects of DNA." At a high-enough level of abstraction, this is true. What would a scientist interested in this level of analysis study? Presumably, they might be interested in addressing questions about the origin of DNA, its stability in various chemical environments, properties of its replication, etc. The Chomskyan tradition, however, attempts to analyze language at this most abstract level while simultaneously rejecting as irrelevant both origins, how languages replicate (i.e., are learned), and their function (i.e., real-world use). One would be forgiven for thinking that the relevance of what remains is hard to grasp.

Judging by the difficulties that linguists have had in constructing even short lists of true linguistic universals (Evans & Levinson, 2009), the differences appear to be substantial. To give just a few examples: some languages appear to have little or no inflectional or derivational affixes (e.g., Vietnamese, Thompson, 1987). Languages vary greatly in the depth of recursion they employ (whether one takes at face value Everett's (2005, 2009) claim that Pirahã lacks recursion entirely, one cannot dismiss the fact that recursion depth differs substantially between languages, e.g., Evans, 2003; Mithun, 2011). Even such apparently fundamental building blocks as nouns, verbs, adjectives, and adverbs are not universal as evidenced by languages such as Straits Salish (Jelinek & Demers, 1994) where the boundaries blur, mirroring a Borgesian fiction:

> …there are no nouns but only impersonal verbs, modified by monosyllabic suffixes or prefixes[s] [F]or example, there is nothing equivalent to our word 'moon', but there is a verb that for us would be 'to moonrise' or 'to moon'. 'The moon rose over the river' would be 'Hlör u fang axaxaxas mlö': … 'Upward, behind the onstreaming, it mooned.'(Borges, 1964, p. 8).

Even in phonology—the part of language perhaps most obviously constrained by physical limitations on production and perception—there are substantial differences in phoneme inventory size, syllable complexity, stress patterns, etc. (as an example, see WALS Chapter 1, Maddieson, Bhattacharya, Smith, & Croft, 2011 to get a sense of differences in consonant inventories and their world-wide distributions). To be sure, there are numerous constraints on cross-linguistic phonological variation. Though here, too, the focus traditionally has been on phonology-internal factors rather than on understanding precisely how vocal production and speech perception shape phonological systems, or understanding the constraints that different environments may place on phonology as a function of e.g., properties of sound transmission through various mediums (see Ember & Ember, 2007 for some intriguing observations and speculations). For example, is it simply a coincidence that "whistled languages" such as Silbo Gomero (e.g., Meyer, 2004) tend to occur in environments that call for a way to communicate across large or difficult to traverse areas? Or do such phonological systems comprise an adaptation to the environment, a solution to a particular problem?

In summary, despite the fact that all languages have certain design features in common, research into these features has been tied to the evolution of language—a topic largely ignored by researchers who take a generative stance on language. At a less abstract level of analysis—one that involves studying the specific structural features of language—analyses of languages have failed to find support for absolute universals. The rule seems to be constrained diversity, not universality.

3.1. Simulating the role of drift and selection pressures in linguistic diversity

Before we present further evidence for this perspective, we showcase a very simple simulation that implements the basic ideas just discussed. So far we have argued that drift and selection, together, bring about diverse linguistic patterns. In fact, selection cannot act without an appropriate amount of variation in a pool of speakers from which particular forms may be adaptive, taken up by other speakers, and thereby become more pervasive in a population. To illustrate more directly the role of selection pressures on linguistic diversity, we designed an agent-based simulation intended to serve as a simple existence proof. Agent-based simulations have been used in various contexts to demonstrate that verbalized theories play out in the manner that we expect, in idealized but computationally-implemented simulations. We examine how communication systems change as a function of drift and selection pressure. In this simple existence proof, we find that even a small amount of selective pressure acting on communication systems can drastically impact the amount of observed diversity of these systems.

In our simulations, as in many others, languages (i.e., grammars) are often defined as feature vectors (e.g., Chater, Reali, & Christiansen, 2009) and language change is quantified as changes to the values of these feature vectors (e.g., Nowak, Komarova, & Niyogi, 2001). Here, we defined language grammars as existing on just two dimensions, with each dimension taking on a real value between 0 and 1. Thus, each language L is defined as a two-element feature vector, (f_1,f_2). Each speaker/comprehender (agent, A) was defined as a pair of vectors, one corresponding to a particular value of the two-feature language spoken by that agent, and one corresponding to the agent's physical location in a simulated terrain, defined by a 100x100 square map: $A = \{(x,y), (f_1,f_2)\}$.

We initialized the simulation by starting 50 agents in the center (location $x = 50$, $y = 50$) and then diffuse according to a set of migration rules. All agents at the beginning spoke the same language, $L_{original} = (.5,.5)$, in accord with the assumption of monogenesis of human language.

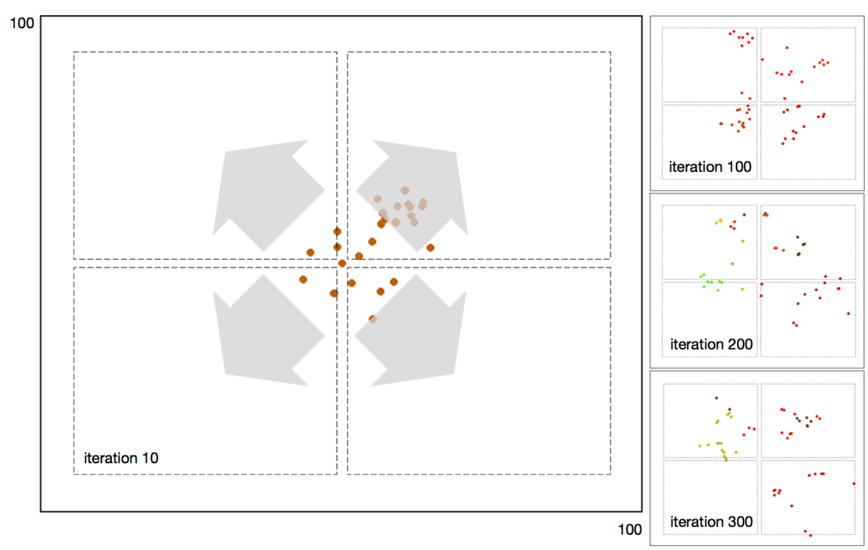

Figure 1: Left panel: A 100x100 grid traverses by 50 agents. Early in a simulation run (e.g., iteration 10), the agents are still near their origin, and their languages are relatively similar. Right panels: As the simulation proceeds, languages drift apart. Note on the four 2x2 dotted quadrants in each terrain: These grids simply illustrate the selection pressures we placed on languages in the second simulation runs described below. The top-right portion of the 100x100 terrain, for example, favored drift towards higher f_1 values; the lower right gridded region favored lower values of f_2. The other two regions were given the remaining selection possibilities (low f_1 value, high f_2 value). Note on color: This is only for visualization. The color of agents is determined by their "language" features, which we use to reveal the presence of agent "dialects" that have drifted apart.

On each iteration, we selected a random set of at most 5 agents that were within 10 units of each other and randomly moved them in any direction on the grid (maximum ± 20 steps). In addition, agents could communicate provided they could "understand" one another. Agents were deemed to understand one other as long as their languages differed by less than the "talk threshold"—the Euclidean distance between their language vectors, here set to $\sqrt{2}/6$ (i.e., 17% of maximum distance; in reality, of course "understanding" is not an all-or-none phenomenon). Finally, we also used language differences to decide migration patterns. We added the constraint such that agents only migrate *together* if they have languages within a distance of $\sqrt{2}/3$ (i.e., 33% of the maximum linguistic difference). This implements the idea that agents in the same region of the landscape who speak the same language are a social group.

We also implemented a notion of linguistic "conformity"—talk like the others talk (Keller, 1994). To this end, agents changed their language to be more similar to each other when they spoke. To implement this, each time agents communicated, they shifted their languages towards their the mean language between them. Finally, drift was implemented as the proportion of the unit space (0-1) that an agent could shift its language up or down on each turn. As an example of the kinds of small changes or tweaks such drift corresponds to, consider the choice of using "whom" vs. "who" in the accusative, or the choice between the prescriptively correct "between you and I" vs. the colloquial "between you and me." The magnitude of the drift parameter controlled the freedom the agents had to "play" with language.

We ran the simulation for 500 iterations using 50 agents and explored a range of drift and selection parameters. At each iteration, one group of agents (maximum $N = 5$) was permitted to migrate on the 100x100 terrain. Also at each generation, all agents were permitted to "communicate" to a group of agents (maximum 10) that were within 10-unit distance around it, that had sufficiently similar languages (as described above).

The results from simulations varying the amount of drift (k, using drift value of $\pm kU(0,1)$) and selection-pressure are shown in Figure 2. Not surprisingly, drift has a large impact on language stability. With too small a drift (essentially perfect language transmission from one generation to the next), the language fluctuates around its initial state of (.5, .5). When drift is increased to 5%, languages become wildly unstable, oscillating radically from one time step to the next (a situation that would prohibit effective communication). With an intermediate amount of drift (1%-3%), the languages diversify while maintaining some stability for a time.

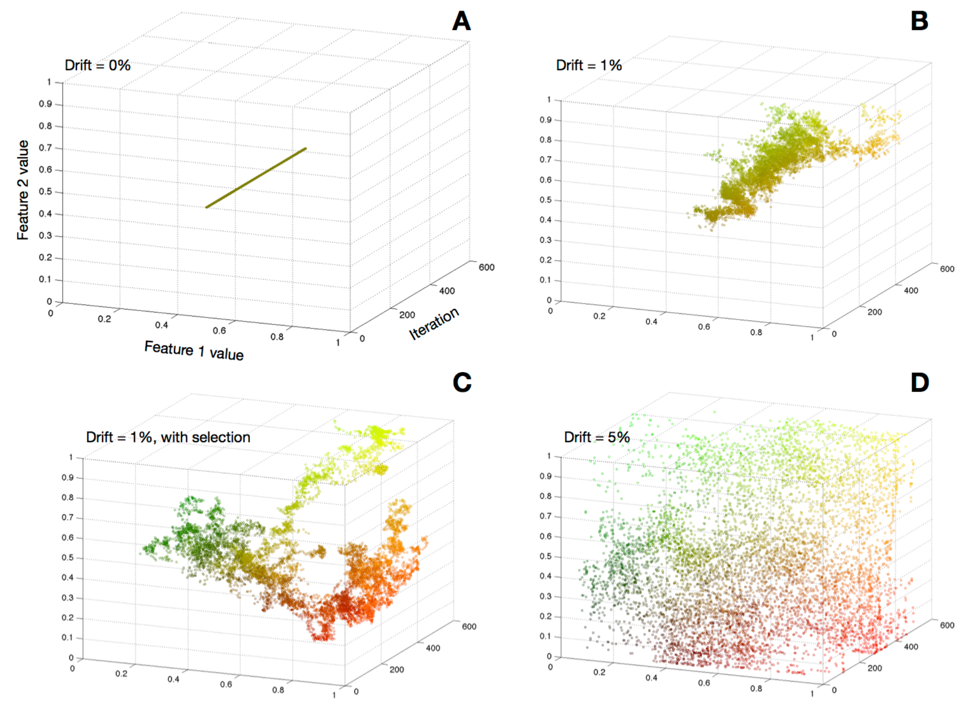

Figure 2: Example runs of the simulation under different parameter values. Colors are only for illustration, and are coded using the feature values, facilitating the observation of dialect formation over time (iterations). See text for details.

To examine the effects selection pressure on linguistic diversity, we implemented a selection pressure by preferentially weighing the agents whose language happened to be most adaptive to the environment in which it happened to find itself. To simulate different environments, we divided the 100x100 grid into four quadrants. In each quadrant, languages with particular feature-values were "favored". For example, in quadrant 1 (top right), languages which happen to have high values on feature 1 would be favored, with no selection pressure applied to feature 2. The selection of particular language-variants was done by increasing the likelihood of agents imitating those using a more adaptive feature than a maladaptive or neutral one. Although we use the word "imitation," the process is not necessarily goal-driven. Imagine an environment in which certain phonemes are poorly transmitted (e.g., high-frequency sounds in humid environments). Individuals using those phonemes, all other things being equal, would be less likely to be imitated than those using alternate forms that are better transmitted.

The effects of selection can be seen in Figures 2 and 3. In Figure 2, each point represents the language state of an agent A and its (f_1, f_2) vector, across time (iterations). Figure 2A shows that, when the simulation has no drift, the two language features stay fixed on their initial values. However, when some drift is added (a 1% perturbation during each interaction, Figure 2B), the languages can begin to explore the parameter space, and come to form agent "dialects." With too high a drift value (5% perturbation, Figure 2D), dialects cannot readily stabilize and languages fluctuate rapidly from iteration to iteration. Things change substantially when the "environment" creates diverse selection pressures. When there is selection pressure, the languages much more rapidly diversity into dialects, even with minimal drift (Figure 2C).

Holding drift rate constant, even a small selection pressure allowed the space of possibilities to be explored more quickly, with the "languages" slowly converging on patterns better adapted to particular regions of the grid. In short, allowing languages to adapt to the environment rather than just change as a function of drift has a profound effect on the level and type of linguistic diversity. On this account, patterns of linguistic diversity are explicable not only in terms of shared history and common descent, but in terms of environmental pressures: Languages spoken in similar social and ecological environments may become more similar as they adapt to common pressures. Figure 3 shows that only a little bit of selection may be necessary for this effect to compound over time. So even if this environmental adaptiveness has a small effect in the overall scheme of language change, it is likely to have important long-term effects on linguistic diversity.

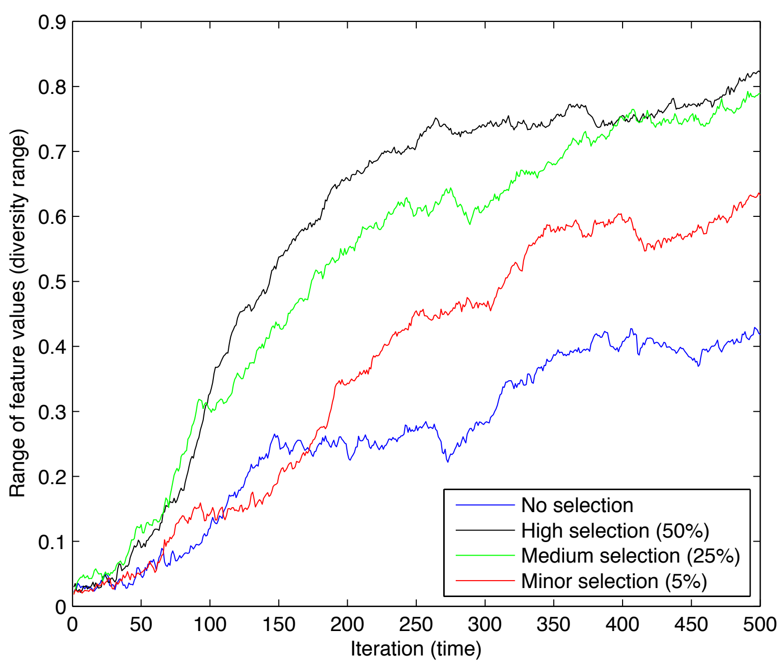

Figure 3: The results of four different runs using 1% drift with varying selection rates. The y-axis shows the average range in f_1 and f_2 across 5 runs of the simulation. Higher scores indicate greater linguistic diversity.

4. The role of drift and selection in explaining linguistic diversity

Our simulation shows how linguistic diversity can arise when drift combines with even a pinch of selection. Although extremely idealized, we view this simulation as a starting point for exploring questions concerning the sources for linguistic diversity.

Once we admit that languages differ, the next step is to try to account for these differences. That is, granting that languages are different, why are the differences what they are? Consider the approach a biologist might take when presented with, for example, differences in beak shape between species of finches. After describing and quantifying the variation, the next step is to understand what factors may be responsible for the observed differences. An obvious place to look in the case of beak shape would be diet and availability of food sources that are more easily or more difficult to access using various beak shapes. Once the associations between beaks and diet are known one can look at how changes in availability of food impact the mortality and reproduction rates of individuals with varying beak shapes across and within a species—direct evidence of a selective pressure on beak shape. As we elaborate in more detail below, we believe that substantial progress in understanding linguistic variability can be made by applying an analogous approach and treating different languages as adaptations to different environments.[3]

The idea that there may be some systematic relationship between language and aspects of the environment, particularly the social, cultural, and technological aspects of the environment, is not a new one. In fact, speculations on the connections between particular grammars and culture appear to be (or at least have been) so commonplace (see Perkins, 1992 for discussion), that in his 1921 book, Sapir admonished all attempts to link language types to what he referred to as "national temperament":

> [A]ll attempts to connect particular types of linguistic morphology with certain correlated stages of cultural development are vain. Rightly understood, such correlations are rubbish. [Both] simple and complex types of language of an indefinite number of varieties may be found spoken at any desired level of cultural advance. When it comes to linguistic form, Plato walks with the Macedonian swineherd, Confucius with the head-hunting savage of Assam. (p. 234)
>
> It is difficult to see what particular causal relations may be expected to subsist between a selected inventory of experience [and] the particular manner in which the society expresses all experience. (p. 233)

[3] It is important to note that just as in biology, there is no requirement that every observed trait should be functional or predictable from some aspect of the environment. Indeed, conditional universals of the form "If a language has property A, it most likely has property B," are a prime example of how a selective force acting on property A may produce property B as a spandrel (that may subsequently become exapted for other functions).

At the same time, Sapir also noticed that language changes were not random, but exhibited what he referred to as the "drift to the invariable word," noting for example that "striving for a simple, unnuanced correspondence between idea and word that [is] very strong in English." (p. 180) Sapir believed that these changes were due to forces internal to language: "Language moves down time in a current of its own making. It has a drift" (p. 160) and that while the lexicon of a language is naturally shaped by the needs of its speakers, "its line of variation, its drift, runs inexorably in the channel ordained for it by its historic antecedents." (p. 232)

The apparent directionality of language change was also described by Jespersen, who made similar observations of language apparently tending to become, over time, more analytic, but unlike Sapir, Jespersen (1922) saw in these changes, a kind of progress: "[There is a] progressive tendency from inseparable irregular conglomerations to freely and regularly combinable short elements," arguing that in "modern" languages, words are shorter, "thus involving less muscular exertion and requiring less time for their enunciation", their formation (i.e., morphology) and syntactic use (i.e., recombination) "present fewer irregularities" and "The clumsy repetitions known under the name of concord have become superfluous" (p. 364).

As we shall see (as noted by Trudgill in this volume, and Trudgill, 1988, 1989, 1993, 2001a; as well as by Christiansen & Chater, 2008; Dahl, 2004; McWhorter, 2001a; Nettle, 1996, 1998a, 1998b; Perkins, 1992; Wray & Grace, 2007), there really is something to this observation. But in ascribing progress to these apparently directional language changes, Jespersen makes the same mistake as someone who, on observing the apparent advantage of the giraffe's long neck, concludes that zebras, antelopes, and the decidedly short-necked gnus, are all at different stages of progress toward girraffean necks. The correct analysis, of course, is that long necks are an adaptation to a particular environment (a niche). Just as we can explain the emergence of and change to physical traits as responses to selective pressures from the environment, we can conceive of culturally-transmitted traits (of which language is only one) as reflecting adaptations to a particular niche. The philosopher Ernst Cassirer expresses a similar idea when he writes that:

> Every classification is directed and dictated by special needs, and it is clear that these needs vary according to the different conditions of man's social and cultural life … Languages vary with the functions they fulfill in the cultures in which they are spoken. (Cassirer, 1945/1962, p. 136).

But what are these "special needs"? One obvious example derives from the referential function of language. Many words, concrete nouns in particular, name objects in our environment, and insofar as there are cross-cultural differences in what needs to be named, the lexicon adapts accordingly. But what about grammatical factors such as verb agreement, cases, and other features that apparently serve purely linguistic functions? What "special need" might these fulfill and what possible conditions of "man's social and cultural life" might vary to as to make some of these linguistic features more or less adaptive?

This is not an easy question. In the biological domain, we have a substantial knowledge base, compiled through observation and theorizing, about at least some functions of various traits. We see birds using their beaks for eating, and so we make the reasonable assumption that differences in beak shape may have something to do with obtaining food. We observe leopards stalk prey and theorize that their coat markings are an adaptation to avoid detection by prey. In inquiring about the functional significance of specific linguistic features, we know far less. What are inflectional evidentials *for*? Person agreement? Complex hierarchies of demonstratives?[4]

Rather than focusing narrowly on explaining why some languages have e.g., complex person agreement, while others do not, one can ask whether particular *types* of languages are more likely to be found in one environment or another. What aspects of environment, of Cassirer's "social and cultural life" are the important ones? Might it matter, for example, if a language is spoken by a thousand versus a million speakers? In an artifact-rich or largely natural environment? In a "society of intimates" or a "society of strangers"? If it borders many languages or is geographically isolated?

5. The importance of learning mechanisms

A useful starting place is the critical but often overlooked observation that languages need to be learnable (Christiansen & Chater, 2008; Deacon, 1997). By definition, an unlearnable language cannot exist. But while all natural languages are constrained by what can be learned by infants, only some languages are additionally constrained by what can be learned by adults. Insofar as children and adults differ in the kinds of linguistic devices they can learn, an immediate prediction is that languages with more nonnative speakers (spoken in what could be called the "exoteric" niche [Thurston, 1989; Wray & Grace, 2007], analogous to Trudgill's use of the term "societies of strangers"; see this volume) will come to have simpler morphological paradigms. Trudgill articulated a version of this hypothesis in perhaps the clearest way:

> Just as complexity increases through time, and survives as the result of the amazing language learning abilities of the human child, so complexity disappears as a result of the lousy language-learning abilities of the human adult. Adult language contact means adult language learning and adult language learning means simplification, most obviously

[4] A broad objection on the grounds that it is impossible to explain language functionally because language is not functional and somehow perfect (Brody, 1998; Lasnik, 2002; Piatelli-Palmarini & Uriagereka, 2004) makes little sense to us and we cannot think of any other domain in which an analogous proposition would be seriously entertained. However, given how little we still know about the functional role of specific features—we are only now starting to systematically catalog and quantify linguistic variation on a large scale (Dryer & Haspelmath, 2011)—it may be premature to theorize from any specific feature and a more productive approach may be one that focuses on broader patterns.

> manifested in a loss of redundancy and irregularity and an increase in transparency. (Trudgill, 2001a, p. 372).

Similar arguments, focusing on the role of the language population on morphological complexity, have also been discussed by McWhorter (2001b, 2002; 2007), Wray and Grace (2007), and a number of contributors to Sampson, Gil, and Trudgill's excellent volume (2009).

A strong test of this hypothesis on a large scale, however, only became possible with the publication of large corpora of grammatical features (e.g., the World Atlas of Language Structures; WALS, Dryer & Haspelmath, 2011), allowing us to examine whether morphological complexity is actually predicted by factors related to exotericity, namely the number of speakers. In a 2010 study, we showed that simply knowing how many people speak a given language, or how widely a language is spoken around the world (in km^2), we could predict, sometimes with very high certainty, some of its grammatical structure. For example, we found that languages spoken by more people, tended to: (1) be less synthetic or fusional, and more analytical in their overall structure, (2) have less complex noun and verb agreement, (3) have simpler overall verb morphology, (4) have fewer noun cases, (5) lack inflectional evidentials, future tense, and aspect markers. Population (as well as geographic spread and number of bordering languages)—the three proxy factors we used to quantify exotericity—predicted over 20 grammatical factors related to morphology (controlling for language family and geography and using Monte Carlo analyses to deal with Galton's problem of non-independent sampling). Overall, the results showed that given the choice of expressing a certain semantic distinction using morphological or lexical means, exotericity was positively correlated with lexical strategies and negatively correlated with morphological encoding of these distinctions. We framed the results in terms of the *Linguistic Niche Hypothesis* (Lupyan & Dale, 2010), arguing that they are indicative of languages evolving to fit the learning constraints of their learners. As a language spreads more widely, and is learned by more adult nonnative speakers, its morphological structure tends to simplify. This is the very process Trudgill envisioned when he wrote:

> Adults [learners]… necessarily subject new languages that they are learning to the process of pidginization… an increase in transparency, by which is meant an increase in forms such as eye-doctor as opposed to optician, and did instead of went. Imperfect learning, that is, leads to the removal of irregular and non-transparent forms which naturally cause problems of memory load for adult learners, and to loss of redundant features. This can in turn lead to an often dramatic increase in analytic over synthetic structures. (Trudgill, 2001b, p. 66)

Our 2010 findings, across over 2,000 languages, suggest that such a process is actually at work. An interesting further test of the hypothesis that exoteric environments, particularly adult learning, produce increases in transparency is a comparison of American and British English. Although nominally the same language, American English is used in a considerably more exoteric setting, as measured by, for instance, the relative proportions of nonnative speakers.[5] According to Ethnologue, about 20% in the US are nonnative speakers versus about 5% in the UK, though the latter number is rapidly increasing. On this niche hypothesis, American English should show a preference for more regular/transparent forms. Consider Trudgill's own example of the lexical form "eye-doctor" compared to the derivational form "optician." A Google search comparing American versus British uses of the terms shows them to have comparable frequencies in webpages originating in the UK: eye-doctor = 2.64 million, optician = 2.95 million; a lexical/morphological ratio of .89. That same analysis performed on US web-pages shows a very different pattern: eye-doctor = 65.6 million, optician = 7.88 million, a lexical/morphological ratio of 8.32. One may wonder if such differences are a symptom of British English being simply more conservative (perhaps owing in part to the smaller speaking population in comparison to American English). Optician, after all, is the original term for a eye-doctor, and so perhaps British English just has a greater bias to hold onto older forms, regardless of what those forms are. Such an account, however, could not explain why British English has apparently been *more* willing to replace the older past-tense form "lighted' with the morphologically irregular "lit" (lighted = 1.64 million, lit = 18.9 million[6], regularity ratio = .087) compared to American English which has retained "lighted" to a greater degree (lighted = 39.9 million, lit = 226 million, regularity ratio = .177; see Dale and Lupyan (2012) for discussion and additional analyses). A common explanation for both patterns is that American English has a stronger affinity for simpler morphology and greater form-to-meaning transparency.

This greater affinity for transparency in American English can also be observed in its greater productivity of certain derivational suffixes. For example, one can refer to retired people as "retired people" or as "retirees." In this case, British English shows a strong preference for the lexical form: "retired people" = 1.66 million, retiree = 396,000, lexical/morphological ratio = 4.19), while American English shows the opposite pattern: "retired people" = 978,000, "retirees" = 37.1 million, lexical/morphological ratio = .026.[7] This pattern appears initially to contradict the earlier observation that American English "prefers" lexical over derived forms (e.g., "eye-doctor" over "optician"). We can resolve this apparent contradiction by considering the difference in the transparency of the form-to-meaning mapping in the two cases. The meaning of retiree, like the meaning of addressee, lessee, interviewee, and payee, is a fairly transparent composition of "one who is" and the meaning of the stem. Thus, retirees are literally, those who are retired. In comparison, the meaning of optician is more opaque. Not only is its medical sense

---

[5] It may be objected that it was British rather than American English that was spread around the world in colonial times. Indeed this is so, but its learning by nonnative speakers and hence the changes we hypothesize were caused to it by this learning, was largely outside the boundaries of England proper.

[6] The number excludes pages that appear to be using lit as an abbreviation for literature.

[7] Interestingly, the popularity of "retiree" in British English appears to be experiencing the kind of rapid increase over the past 15 years that it experienced in American English in the 1950s, according to a Google Ngram analysis.

not obviously derivable from the stem "optic," but the use of suffix –ian to mean (roughly) "one who performs [stem] as a profession" is in competition with other suffixes with similar meanings, cf. electric/electrician, pediatrics/pediatrician, but neurology/neurologist, surgery/surgeon, plumbing/plumber.

Analyses such as the ones above, although suggestive, are clearly preliminary. Much more rigorous work is needed to determine the degree to which there is a systematic bias in American English for more transparent form-to-meaning mapping, and the degree to which this is true more generally. For instance, one might make the opposite prediction for Quebecois French as compared to the French in the home country insofar as Quebecois became more insulated from influences of outsiders. Beyond the specifics, however, the larger claim we are making is that differences between American and British English, can be understood in part as the consequence the two Englishes adapting to slightly difference niches.

If one assumes that children are better learners of opaque form-to-meaning mappings than adults, it is easy to see how nonnative speakers in a language act as a kind of bottleneck. But what may be less obvious is how the presence of nonnative speakers impacts the native-speaking population. To understand how this works, we need to recall that languages need to be learned by their speakers (we have purposely avoided the term "acquire" commonly used when referring to language learning in infancy). The exact form of a language a child will learn depends strongly on the input. If the input is provided, even in small part, by nonnative speakers, that may affect the language the child goes on to learn. For example, in a survey of 188 individuals in Senegal who listed Bambara as their native language, Bambara was the father's native language in 16%, the mother's in 19%, the native language of both parents in 26%, and the native language of neither parent in 39% (Calvet, 2006). Although children are learning Bambara from a young age and are, in theory, fully capable of learning whatever morphology it possesses, in such a multilingual environment much of the Bambara they hear would come from nonnative speakers. Thus, whatever aspects of Bambara were difficult for the parents to learn are more likely to be passed on to the offspring in a revised form.[8]

To investigate further the influence of even a small amount of exposure to nonnative speakers, Dale and Lupyan (2012) elicited acceptability ratings of overregularized sentences such as "He speeded down the road" and "They sneaked around" from 95 native American English speakers from around the country. The results showed that the degree of acceptability of such sentences (partialing out several factors like level of education) was predicted by the amount of childhood exposure to nonnative English speakers (derived from self-report and US Census records based on the proportion of nonnative speakers in the US state where they grew up). In the same paper, we describe a series of agent-based simulations that show how even a

[8] We do not mean to suggest that infants simply copy what they hear. All language learners generalize beyond their input. But if, for example, a particular morphological distinction is simply absent from the input, then it is unlikely that the learner is going to reinvent it on their own. If they did, languages would not be as different as they are.

small bias against complex morphology can impact the level of morphological specification that a language comes to possess.

6. The child-adult learnability tradeoff

We have thus far argued that while all languages are necessarily constrained by what can be learned by infants, but only some—the languages occupying the more exoteric niche—are further constrained by the limitations of adult language learning (though what precisely those limitations are, is still very poorly understood). But why does complex morphology arise in the first place? It has often been noted that languages are far more complex than they need to be (e.g., Premack, 1986), and as Gil (2009) convincingly argues, all this extra complexity does not seem necessary given how much can be accomplished with languages lack most of these "baroque accretions" (see McWhorter, 2001b for discussion). From a linguistic-niche perspective, one possible answer to this puzzle is that complex surface morphology and paradigms that present difficulties for the adult learner actually *benefit* child learners. Consider, for example, agreement, what Jespersen referred to as "clumsy repetitions known under the name of concord." But it is precisely the repetition, and the overall increased redundancy and overspecification of languages spoken in the esoteric niche that may provide learning benefits to children. While agreement can pose challenges for adult L2 learners, agreement (as well as grammatical gender, complex demonstratives, morphologically encoded aspect, evidentiality, etc.) may *facilitate* language learning by children by providing them with additional cues helping to ground the linguistic stream to the goings-on in the environment. The intuition for this admittedly counterintuitive claim is that in comparison to adults who can deploy powerful pragmatics, theory-of-mind, and general world knowledge to make sense of ambiguous utterances, children do not yet have these mechanisms at their disposal. Thus, encoding aspect, gender, evidentiality, etc. grammatically (with its corresponding increase in redundancy), may baffle the adult, but be beneficial to the child learner.[9]

An immediate objection to the idea that richly inflected languages are better adapted to child learners is that it seems to suggest that morphologically complex (and more opaque) languages ought to be learned better by children compared to morphologically simple languages such as English. There is indeed evidence of differences in learning rates across languages (e.g., Slobin & Bever, 1982), and some evidence of faster learning by children of more complex inflectional systems. For example, Devescovi et al. (2005) observed that Italian children require fewer words to extrapolate grammatical regularities of Italian compared to children learning English, a difference the authors ascribed to the richer inflectional system of Italian which provides the children with increased learning opportunities.

---

[9] As a demonstration that languages spoken in more esoteric niches are indeed more redundant, Lupyan and Dale (2010) quantified informational redundancy in terms of the Hoffman codes (approximated by zipping a text file; redundancy is proportional to the degree to which the file can be compressed). We found that, indeed, languages spoken by fewer people (typically those with few nonnative speakers) were considerably more compressible, i.e., had greater redundancy.

Despite some evidence of cross-linguistic differences in rates of language learning, children learning morphologically richer and more overspecified languages do not appear to learn them substantially faster. Cross-linguistic differences in language learning appear to be fairly minor. Where there *are* large cross-linguistic differences, however, is in the amount of language directed at prelinguistic children (e.g., Johnston & Wong, 2002; Richman, Miller, & LeVine, 1992; Tamis-LeMonda, Song, Leavell, Kahana-Kalman, & Yoshikawa, 2012). An intriguing possibility is that such differences interact with the grammar of the language being learned by the children. If more richly inflected (and hence more redundant) languages are especially well adapted for child learning, perhaps they can be learned with less input. As a language becomes exposed to the learning constraints of adults and loses some of the inflectional richness (and with it, redundancy), children require more input to learn it. There is now considerable evidence showing just how sensitive English-learning children are to reduction in input (Hart & Risley, 1995; Hoff, 2003; Huttenlocher, Vasilyeva, Cymerman, & Levine, 2002). Direct comparisons of input sensitivity between languages, however, are lacking at present.

There are at least two additional domains in which languages may adapt to child learners. First, there may be tradeoffs between morphological and syntactic complexity. There is clear evidence that processing sentences with deeper embeddings requires greater working memory (Lewis, 1996) which, in the case of young children, is in relatively short supply (e.g., Gathercole, Pickering, Ambridge, & Wearing, 2004). Insofar as morphologically complexity tends to allow for simpler syntax (particularly in the case of syntactic embedding, Evans & Levinson, 2009), one can ask whether these simpler embeddings made possible by richer inflectional systems facilitate language learning in children.

A second domain is in how inflections are implemented, such as whether a language is predominantly suffixing or prefixing We have painted inflectional systems in very broad strokes, speaking of richly inflected languages, or ones with little inflection, but of course there is substantial variability in the form those inflections take. One difference is whether the inflections take the form of suffixes or prefixes. According to WALS, there are far more languages that are strongly biased for inflectional suffixing (382 languages from 66 language families) compared to prefixing (54 languages from 14 language families). Based on these data, one may conclude that suffixation is, in some way, more natural. But if we now look at the demographics of the languages that use suffixation versus prefixation, we get a very different picture. In an analysis we reported earlier (Lupyan & Dale, 2010, supplementary materials) we found that languages that primarily use prefixation have, on average, 10 times greater population than those that primarily use suffixation.[10] In other words, languages we argue to be most adapted for child

[10] A similar mismatch between "naturalness" according to number of languages/language families demonstrating a given trait, according to demographics, was also notably found for basic word order. Despite SOV being the most widespread word-order, SVO is far more common in languages with larger speaking populations.

learning favor suffixes, while those that have been shaped by adult learning, favor prefixes.[11] We would therefore predict that suffixing is an adaptation to esoteric niches. Indeed, there is some evidence that suffixes are easier to learn for infants than prefixes (Kuczaj, 1979; Slobin, 1979, 1985), and there is some indication from experimental studies that prefixes are easier to learn for adults compared to suffixes (Frigo & McDonald, 1998; MacWhinney, 1983; St. Clair, Monaghan, & Ramscar, 2009).

In summary, although there now appears to be converging evidence for the connection between adult language learning and morphological simplification, the reasons why languages have complex morphological systems to begin with are more puzzling. We have argued here that rather than comprising "baroque accretions," complex morphological systems may play a role in facilitating language learning by children and thus comprise an adaptation to the esoteric niche. We admit, however, that this proposal remains tentative and awaits more rigorous empirical tests.

7. Ecological constraints on language structures

In discussing the environment to which languages adapt we have focused on social and demographic factors, such as the effect of a language being constrained by child leaners or a combination of child and adult learners, and languages being used by societies of intimates versus societies of strangers. We have said little about the ways in which grammars may adapt to exogenous factors such as the physical environment in which the language is learned and used. Here, we consider several examples from the domain of spatial language.

Although all languages have ways of expressing the relative locations of objects or people, the precise means of doing so differs. One source of such difference lies in the system of demonstratives, terms such as this, that, here, and there. In some languages like English, the demonstrative systems are quite sparse and underdetermined. To make sense of an expression such as "I am here," one needs to know quite precisely the context of the utterance. Does the speaker mean *here* in the city? *here* in the building? *here* in the office, *here* at the restaurant? Of course, one can optionally add this information, but, nothing about the word "here" specifies where the "here" is. Such systems contrast with those that require speakers to encode relative location much more precisely using demonstratives and other devices such as deictic adverbs (e.g., Denny, 1978, 1982; McWhorter, 2002 for a discussion of English as compared to other Germanic languages). To what degree may such differences reflect adaptations to different environments? Denny (1978) proposed that certain spatial systems seem particularly well-suited for describing relative locations in artifact-sparse environments, in which the familiar English system of demonstratives and deictic adverbs, would appear to fail. In English, we regularly refer to regions of space with phrases such as "That one across the street" or "To the left of that mailbox." But such expressions would be of limited use in an environment in which no such

[11] One may wonder why, given the exotericity of numerous Indo-European languages, none have strong inflectional prefixing. Our uninformed guess is that inflectional prefixes were outside the variation of languages and thus could not be selected.

reference points exist. One solution is to center the space on speakers and listeners instead. As Denny writes:

> [In] a natural environment of non-human spaces one way to relate space to human activity is to use deictic spatial concepts, to center space on the speaker or other participants. [In a man-made environment this is less necessary… we can use non-deictic locatives (down the road, around the corner) which will relate space to human acts quite directly since the places mentioned are all artifacts designed to aid such acts (1978, p. 80).

Is the presence of complex demonstrative systems in some languages simply a coincidence? Or, might they be seen as an example of a linguistic adaptations to particular ecology? No one, to our knowledge, has looked at relationships between ecologies and language structures. In a feasibility study, we used the Standard Cross Cultural Sample (SCCS, White, 2007), an ethnographic database of 186 cultures, to test the generality of Denny's (1978) observation inspired by the study of spatial terms in Eastern Eskimo. Without the ability to say things like "next to the mailbox," the language is, on the present account, under a selective pressure to develop complex speaker- and listener-centered spatial terms (that are unnecessary in an object-rich environment). We combined the biome factor from SCCS (desert, tropics, tundra, etc.) with the measure of spatial-term distinctions from WALS (because SCCS does not include information on specific languages, this analysis was done at the level of language families). Not only did languages spoken in the five biomes differ significantly in the complexity of demonstratives, but as predicted, the languages spoken in the most sparse biome (tundra and taiga) had demonstrative systems with reliably more remoteness distinctions (the only factor coded by WALS) compared to languages spoken in other biomes. In these results we see hints much larger-scale analyses aiming to understand how particular language structures—morphological, syntactic, semantic, and phonological—can be understood in light of the ecological factors influencing the language.

8. Conclusion

Our main claim in this chapter is that it is impossible to understand why there are so many languages and why languages differ as they do without taking into account the selective pressures that have operated and continue to operate on languages. These pressures can be both endogenous such as cognitive limitations (which may differ quite drastically for child and adult language learners with consequences for languages with many versus few adult learners), and exogenous such as ecological factors in which the language is used.

At present, we would characterize our state of knowledge in understanding what these pressures are, and how they operate, as minimal. However, we foresee tantalizing possibilities in research programs that combine descriptive linguistics datasets with anthropological data, ecological information, literature on child-language, and finally, studies that use artificial-language learning paradigms to study experimentally how languages are influenced by the cognitive constraints of the learner (e.g., Ellefson & Christiansen, 2000; Monaghan, Christiansen, & Fitneva, 2011; St. Clair et al., 2009), and by ecological factors (we have nothing to cite here owing to an absence of any research in this area).

To make progress, however, we must abandon the puzzling dogmas about all languages being equally complex (see Geoffrey Sampson et al., 2009 for discussion), about linguistic variation being irrelevant to understanding the human language capacity, and for language somehow being immune from the functionalism inherent to all other evolutionary traits. Language—both the human capacity for language, and specific grammars—must be viewed as functional systems. The dictum of Dan Slobin makes for an apt conclusion:

> The acquisition and development of any linguistic form or construction must be considered in the light of its "functional load" within the language and speech community (Slobin, 1997, p. 35).

References

Borges, J. L. (1964). *Labyrinths: Selected Stories & Other Writings*. New Directions Publishing.

Brody, M. (1998). The Minimalist Program and a Perfect Syntax: A Critical Notice of Noam Chomsky's The Minimalist Program. *Mind & Language*, *13*(2), 205–214. doi:10.1111/1468-0017.00074

Calvet, L.-J. (2006). *Towards an Ecology of World Languages*. Polity.

Cassirer, E. (1962). *An Essay on Man: An Introduction to a Philosophy of Human Culture*. Yale University Press.

Chater, N., Reali, F., & Christiansen, M. H. (2009). Restrictions on biological adaptation in language evolution. *Proceedings of the National Academy of Sciences*, *106*(4), 1015.

Christiansen, M. H., & Chater, N. (2008). Language as Shaped by the Brain. *Behavioral and Brain Sciences*, *31*(5), 489–509.

Dahl, Ö. (2004). *The Growth and Maintenance of Linguistic Complexity*. John Benjamins Publishing Co.

Dale, R. A., & Lupyan, G. (2012). Understanding the origins of morphological diversity: The linguistic niche hypothesis. *Advances in Complex Systems*, *15*(3), 1150017–1–1150017–16. doi:10.1142/S0219525911500172

Deacon, T. (1997). *The Symbolic Species: The Co-Evolution of Language and the Brain*. London: Allen Lane: The Penguin Press.

Denny, J. P. (1978). Locating the universals in lexical systems for spatial deixis. In D. Darkas, W. M. Jacobsen, & K. W. Todrys (Eds.), *Papers from the Parasessions of the Lexicon* (pp. 71–84). Chicago Linguistic Society.

Denny, J. P. (1982). Semantics of the Inuktitut (Eskimo) Spatial Deictics. *International Journal of American Linguistics*, *48*(4), 359–384.

Devescovi, A., Caselli, M. C., Marchione, D., Pasqualetti, P., Reilly, J., & Bates, E. (2005). A cross linguistic study of the relationship between grammar and lexical development. *Journal of Child Language*, *32*(4), 759–786.

Dryer, M., & Haspelmath, M. (Eds.). (2011). The world atlas of language structures online. *Available at http://wals.info*. Munich: Max Planck Digital Library.

Ellefson, M. R., & Christiansen, M. H. (2000). Subjacency constraints without universal grammar: Evidence from artificial language learning and connectionist modeling. *The Proceedings of the 22nd Annual Conference of the Cognitive Science Societ* (pp. 645–650). Mahwah, NJ: Lawrence Erlbaum. Retrieved from http://cornell.academia.edu/MortenHChristiansen/Papers/47716/Subjacency_constraints_without_universal_grammar_Evidence_from_artificial_language_learning_and_connectionist_modeling

Ember, C., R., & Ember, M. (2007). Climate, Econiche, and Sexuality: Influences on Sonority in Language. *American Anthropologist*, *109*(1), 180–185. doi:10.1525/aa.2007.109.1.180

Evans, N. (2003). *Bininj Gun-wok : a pan-dialectal grammar of Mayali, Kunwinjku and Kune* (Vols. 1-2). Australian National University. Research School of Pacific and Asian Studies. Retrieved from http://en.scientificcommons.org/48029314

Evans, N., & Levinson, S. C. (2009). The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences*, *32*(05), 429. doi:10.1017/S0140525X0999094X

Frigo, L., & McDonald, J. L. (1998). Properties of Phonological Markers That Affect the Acquisition of Gender-Like Subclasses, ,. *Journal of Memory and Language*, *39*(2), 218–245. doi:10.1006/jmla.1998.2569

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental psychology*, *40*(2), 177–190.

Gil, D. (2009). How Much Grammar Does It Take to Sail a Boat? In G. Sampson, D. Gil, & P. Trudgill (Eds.), *Language Complexity as an Evolving Variable* (pp. 19–33). Oxford University Press.

Gordon, R. G. (2005). *Ethnologue: Languages of the World, 15th Edition* (Fifteenth.). SIL International.

Hockett, C. F. (1966). The problem of universals in language. In J. H. Greenberg (Ed.), *Universals of language (2nd. ed.)* (Vol. 2, pp. 1–29). Cambride, MA: The MIT Press.

Hoff, E. (2003). The Specificity of Environmental Influence: Socioeconomic Status Affects Early Vocabulary Development Via Maternal Speech. *Child Development*, *74*(5), 1368–1378. doi:10.1111/1467-8624.00612

Huttenlocher, J., Vasilyeva, M., Cymerman, E., & Levine, S. (2002). Language input and child syntax. *Cognitive psychology*, *45*(3), 337–374.

Jelinek, E., & Demers, R. A. (1994). Predicates and pronominal arguments in Straits Salish. *Language*, *70*, 697–736.

Jespersen, O. (1922). *Language; its nature, development and origin*. London : G. Allen & Unwin, ltd. Retrieved from http://archive.org/details/languageitsnatur00jespiala

Johnston, J. R., & Wong, M.-Y. A. (2002). Cultural Differences in Beliefs and Practices Concerning Talk to Children. *Journal of Speech, Language, and Hearing Research*, *45*(5), 916–926. doi:10.1044/1092-4388(2002/074)

Keller, R. (1994). *On language change: The invisible hand in language*. Routledge. Retrieved from http://books.google.com/books?hl=en&lr=&id=4ML-4FPf_uYC&oi=fnd&pg=PP1&dq=rudi+keller&ots=BUBeoY7oqp&sig=5GllniIQsuXycNKCDIPpIpKgIe8

Kuczaj, S. A. I. (1979). Evidence for a Language Learning Strategy: On the Relative Ease of Acquisition of Prefixes and Suffixes. *Child Development*, *50*(1), 1–13.

Lasnik, H. (2002). The minimalist program in syntax. *Trends in Cognitive Sciences*, *6*(10), 432–437.

Lewis, R. L. (1996). Interference in short-term memory: the magical number two (or three) in sentence processing. *Journal of psycholinguistic research*, *25*(1), 93–115.

Lupyan, G., & Dale, R. A. (2010). Language Structure Is Partly Determined by Social Structure. *PLoS ONE*, *5*(1), e8559. doi:10.1371/journal.pone.0008559

MacWhinney, B. (1983). Miniature linguistic systems as tests of the use of universal operating principles in second-language learning by children and adults. *Journal of Psycholinguistic Research*, *12*(5), 467–478. doi:10.1007/BF01068027

Maddieson, I., Bhattacharya, T., Smith, D. E., & Croft, W. (2011). Geographical distribution of phonological complexity. *Linguistic Typology*, *15*(2), 267–279.

Mcwhorter, J. (2007). *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars*. Oxford University Press, USA.

McWhorter, J. H. (2001a). The world's simplest grammars are creole grammars. *Linguistic Typology*, *5*(2-3), 125–166.

McWhorter, J. H. (2001b). The world's simplest grammars are creole grammars. *Linguistic Typology*, *5*(2-3), 125–166.

McWhorter, J. H. (2002). What happened to English? *Diachronica*, *19*(2), 217–272.

Meyer, J. (2004). Bioacoustics of human whistled languages: an alternative approach to the cognitive processes of language. *Anais da Academia Brasileira de Ciências*, *76*(2), 406–412. doi:10.1590/S0001-37652004000200033

Mithun, M. (2011). How to avoid subordination. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society*, *10*, 493–509.

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*(3), 325–347. doi:10.1037/a0022924

Nettle, D. (1996). Language Diversity in West Africa: An Ecological Approach. *Journal of Anthropological Archaeology*, *15*(4), 403–438. doi:10.1006/jaar.1996.0015

Nettle, D. (1998a). Coevolution of phonology and the lexicon in twelve languages of West Africa. *Journal of Quantitative Linguistics*, *5*(3), 240–245. doi:10.1080/09296179808590132

Nettle, D. (1998b). Explaining Global Patterns of Language Diversity. *Journal of Anthropological Archaeology*, *17*(4), 354–374. doi:10.1006/jaar.1998.0328

Nowak, M. A., Komarova, N. L., & Niyogi, P. (2001). Evolution of Universal Grammar. *Science*, *291*, 114–118.

Perkins, R. D. (1992). *Deixis, grammar, and culture* (Vol. 24). J. Benjamins Pub. co. Retrieved from http://books.google.com/books?hl=en&lr=&id=JWGK5Zf2TZYC&oi=fnd&pg=PP2&dq=deixis+cultural+complexity&ots=eE0FZ1yIlh&sig=i8u7XB3rOpjm0nv1xFQjmK9wfGs

Piatelli-Palmarini, M., & Uriagereka, J. (2004). The immune syntax: the evolution of the language virus. *Variation and universals in biolinguistics. Oxford: Elsevier*. Retrieved from http://www.biolinguistics.uqam.ca/Piattelli-Palmarini%26Uriagereka_2004.pdf

Pinker, S. (1994). *The Language Instinct*. New York: Harper Collins.

Premack, D. (1986). *Gavagai! or the Future History of the Animal Language Controversy*. Cambridge, MA: MIT Press.

Richman, A. L., Miller, P. M., & LeVine, R. A. (1992). Cultural and educational variations in maternal responsiveness. *Developmental Psychology*, *28*(4), 614.

Sampson, Geoffrey, Gil, D., & Trudgill, P. (2009). *Language Complexity as an Evolving Variable*. Oxford University Press, USA.

Sapir, E. (1921). *Language: An Introduction to the Study of Speech*. Dover Publications.

Slobin, D. . (1979). *Psycholinguistics* (2nd ed.). Scott Foresman & Co.

Slobin, D. I. (1985). Crosslinguistic evidence for the language making capacity. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition. Vol 2: Theoretical Issues* (pp. 1157–1256). Hillsdale, NJ: Erlbaum.

Slobin, D. I. (1997). The Universal, the Typological, and the Particular in Acquisition. In D. I. Slobin (Ed.), *The crosslinguistic study of language acquisition. Vol 5: Expanding the Contexts* (pp. 1–39). Mahwah, NJ: Lawrence Erlbaum Associates. Retrieved from http://ihd.berkeley.edu/Slobin-Language%20Acquisition/(1982)%20Slobin%20-%20Universal%20and%20particular%20in%20language%20acquisition.pdf

Slobin, D. I., & Bever, T. G. (1982). Children use canonical sentence schemas: A crosslinguistic study of word order and inflections. *Cognition*, *12*(3), 229–265. doi:10.1016/0010-0277(82)90033-6

St. Clair, M. C., Monaghan, P., & Ramscar, M. (2009). Relationships Between Language Structure and Language Learning: The Suffixing Preference and Grammatical Categorization. *Cognitive Science*, *33*(7), 1317–1329. doi:10.1111/j.1551-6709.2009.01065.x

Tamis-LeMonda, C. S., Song, L., Leavell, A. S., Kahana-Kalman, R., & Yoshikawa, H. (2012). Ethnic differences in mother-infant language and gestural communications are associated with specific skills in infants. *Developmental Science*, *15*(3), 384–397. doi:10.1111/j.1467-7687.2012.01136.x

Thompson, L. C. (1987). *A Vietnamese reference grammar* (Vol. 13). University of Hawaii Press. Retrieved from http://books.google.com/books?hl=en&lr=&id=XN7SHNeZ_ksC&oi=fnd&pg=PR13&dq=1987+thompson+vietnamese&ots=NCbsLS5lg8&sig=tHC3OH6gwoOC-08JWGblvpfLqmQ

Trudgill, P. (1988). On the role of dialect contact and inter-dialect in linguistic change. In J. Fisiak (Ed.), *Historical dialectology. Regional and Social*, Studies and Monographs 37 (pp. 547–563). Berlin: Walter de Gruyter.

Trudgill, P. (1989). Contact and isolation in linguistic change. In L. E. Breivik & E. H. Jahr (Eds.), *Language change: Contributions to the study of its causes*, Trends in linguistics, studies and monographs (Vol. 43, pp. 227–237). Amsterdam: Mouton de Gruyter.

Trudgill, P. (1993). Dialect typology: Phonological aspects. In G. Aurrekoetxea & X. Videgain (Eds.), *Nazioarteko Dialektologia Blitzarra: Agiriak* (pp. 659–666).

Trudgill, P. (2001a). Contact and simplification: Historical baggage and directionality in linguistic change. *Linguistic Typology*, *5*(2-3), 371–374.

Trudgill, P. (2001b). *Sociolinguistics: An Introduction to Language and Society, Fourth Edition* (4th ed.). Penguin (Non-Classics).

Weiner, J. (1995). *The Beak of the Finch: A Story of Evolution in Our Time* (1st Vintage Bks Ed, Jun. 1995.). Vintage.

White, D. R. (2007). Standard Cross-Cultural Sample Free Distribution Site (UC Irvine). Retrieved March 31, 2009, from http://eclectic.ss.uci.edu/~drwhite/courses/index.html

Wray, A., & Grace, G. W. (2007). The consequences of talking to strangers : Evolutionary corollaries of socio-cultural influences on linguistic form. *Lingua*, *117*(3), 543–578.