PHONEME INVENTORY SIZE DISTRIBUTIONS AND THE ORIGINS OF THE DUALITY OF PATTERNING

LUKE FLEMING

Département d'anthropologie, l'Université de Montréal,
Montréal, Canada
luke.fleming@umontreal.ca

Atkinson 2011 claims that phoneme inventories are largest in Africa and smaller elsewhere, and that this clinal distribution reflects a 'founder-effect' of human migrations 'out-of-Africa'. Because of the way in which velaric ingressive and pulmonic egressive airstream mechanisms combine to create extra-large consonant inventories, click languages have the largest phoneme inventories of all. Critics question why phoneme inventory size, but not other properties of language, should leave a trace of the origin and dispersal of natural language. This paper argues that large phoneme inventories would likely have been characteristic of the first fully modern languages if we assume, following Hockett 1960, that duality of patterning was the last 'design feature' of language to emerge. The diachronic trajectories of sign languages and writing systems illustrate that dually patterned phonologies—where minimal units of linguistic form (or phonemes) capable of distinguishing semantic units (or morphemes) are not meaningful in themselves—are often preceded by a stage in which minimal units of form map directly onto semantic functions. Click articulations would have been essential in elaborating large inventories, and thus large vocabularies, in spoken languages lacking duality of patterning. The contemporary distribution of phonemic clicks offers support for the hypothesis, as genetic studies increasingly point to an eastern or southern African origin for modern humans.

1. Introduction

Atkinson (2011) surveyed phonological data from over 500 languages to argue that "phonemic diversity" is greatest in Africa, and "declines with distance from Africa", as with genetic and phenotypic diversity, and thus "supports an African origin of modern human languages" (Atkinson, 2011:346). Atkinson's theory, drawing on Hay and Bauer (2007), presupposes a constant of phoneme inventory growth and reduction as a function of population size. However, the positive correlation between phoneme inventory size and population size, if valid (cf. Donohue & Nichols, 2011), is a proxy for complex social and linguistic processes (Trudgill, 2011). The population-size-to-phoneme-

inventory-size correlation seems to hold best for complex, agricultural, state-based societies of the kind found only since the Neolithic. Population size effects on phoneme inventories would likely not be have been marked among the Paleolithic hunter-gatherer societies—of small populations, in absolute terms—responsible for human range expansion (Bowern, 2011). Furthermore, evidence for a natural rate of increase of phoneme inventories (as assumed in Perreault and Mathew [2012]) is lacking (Ringe, 2011). In sum, the set of explanatory principles posited in Atkinson (2011) seem ill-suited to explain why the highest levels of genetic diversity and the largest phoneme inventories are both found in Africa.

An even larger global survey of more than 2,000 languages in Creanza et al. (2015) questions whether a "signal" of African origins can even be found in phoneme inventory distributions. Nevertheless, the authors do find a rough correlation between genetically polymorphic populations (the top fifth of which are all found in Africa) and languages with large phoneme inventories (the top 28 languages with the most phonemes in the 2,000+ sample are all from Africa) (Creanza et al., 2015:1267, Table S4). In this paper I suggest an alternative way to look at the intriguing idea that the distribution of phoneme inventories by size reflects human language origins and dispersal.

One way in which to get around the multiple methodological and empirical problems facing the founder effect model (albeit keeping Atkinson's assumption of linguistic monogenesis) is to stipulate that the earliest natural languages had much larger phoneme inventories than contemporary languages. This possibility has likely not been pursued because of its seeming violation of the uniformitarian hypothesis. However, the uniformitarian hypothesis in historical linguistics should imply only that the processes of change affecting prehistoric languages were the same as those affecting contemporary languages, not that the kinds of languages upon which they operated had the same typological profiles as contemporary ones (Comrie, 2003; Newmeyer, 2002). The first modern human languages likely had features derived from the protolanguages out of which they emerged (Nichols, 2011). Since protolanguages were organized by different functional principles than modern languages we might expect the first human languages to be typologically anomalous with respect to contemporary languages, although we cannot know a priori along which axes they differed.

If the first natural languages had extra-large phoneme inventories a much modified version of Atkinson's linguistic-founder effect could still apply, even without the Hay and Bauer population-size-to-phoneme-inventory-size correlations. There is a strong correlation between consonant inventory size and articulatory complexity; the larger a phonemic inventory the more articulatorily complex segments it possesses (Lindblom & Maddieson, 1988). If we assume that the first languages had large inventories we can also assume that they had complex articulations. In this scenario, as marked phonological features (e.g. ejectives, implosives, clicks, etc.) were lost in daughter languages there would have been low probability of the renewal of those features outside of contact

areas where they were present and could be renovated via horizontal cultural transmission.

This hypothesis does a decent job of accounting for the observed fit between largest phoneme inventories and most genetically diverse populations. The rule of thumb in historical linguistics is that the homeland of a language family is most likely "the area represented by the greatest diversity (largest number of subgroups) for which the minimum number of moves (migrations) would be required to bring the speakers of the diverse languages back to one place" (Campbell, 2004:430). On the assumption of extra-large phoneme inventories as the initial condition of spoken language, a good guess for the homeland of language would be a site where languages affiliated with distinct language families, all possessing extra-large phoneme inventories of high articulatory complexity, are found clustered together. Importantly, the largest phonemic inventories in the world are found in the areally linked, but genealogically separate, Kx'a, Tuu, and Khoe language families (cf. Greenberg's "Khoisan") (see section 4). And these "core click languages" (Güldemann & Stoneking, 2008) are spoken in southern Africa, in the same geographical area where some of the most genetically polymorphous populations are found (Knight, 2003; Tishkoff et al., 2007).

If this hypothesis offers a plausible cause for the correlation between largest phoneme inventories and most genetically diverse populations, it is still unclear what reason could motivate us to accept the counterintuitive conclusion that the earliest human languages had markedly larger and more complex phonological inventories than contemporary ones. In the following sections I argue that the evolution of the phonology-semantics interface may offer such a rationale.

2. Duality of Patterning and the Lexicon Design Problem

As Hockett ([1960] and Martinet [1960]) observed, languages are patterned at two distinct levels or planes. At one level of patterning meaningful units of language, words and other morphemes, combine with one another to form phrases and sentences. But those morphemes are themselves made up of another order of units, phonemes. The phonemes of spoken languages are sound segments capable of distinguishing meaningful units from one another, but not inherently meaningful in and of themselves. The organization of human language thus relies upon the combination of meaningless phonemic units to yield meaningful morphemic units, which themselves combine with one another to form sentences—two levels of patterning.

Duality of patterning is often seen as essential for the development of a large lexicon, and then most pressingly in the spoken modality of language (de Boer, Sandler, & Kirby, 2012). Since the need for duality of patterning only arises under conditions where a large number of meaningful speech signals are already employed, Hockett (1960:96) reasoned that it would be one of the last developments in the natural history of language. In a spoken language lacking

duality of patterning, minimal units of segmentable form would map directly onto semantic functions. Following Hjelmslev (1961:113), I call such linguistic systems MONOPLANAR. Since lexicons of languages of this type are limited by the inventory of representational forms which they employ, they tend towards the production a large number of perceptibly different form types (see section 3). However, to represent each morpheme in the lexicon, especially in the oral-aural modality, by means of a unique sign-form would be extremely burdensome. It may be true that visually based codes, like sign and writing, allow for indefinitely large monoplanar lexicons. For the oral-aural modality, however, there are rather stringent limits on the number of phonetic distinctions that humans can perceptually make. Duality of patterning solves this problem by adding a level of complexity in the form of a mediating level between representational-form and semantic-meaning. The emergence of duality of patterning in sign languages and writing systems offers good evidence for these claims.

3. Diachronies of Duality of Patterning Beyond the Spoken Modality

There are concrete obstacles to reconstructing the emergence of the duality of patterning in speech. The dominance of speech, among linguistic modalities, ironically means that we have perhaps the most limited knowledge about the emergence, ex nihilo, of spoken languages. Because of their polygenetic development, visually based codes offer the best empirical data for studying the emergence of the duality of patterning.

Sign languages, like spoken languages, typically exhibit duality of patterning in their phonological structures (Stokoe, 1960). Phonological categories typically organized into distinctive features in sign languages are Hand Configuration, Location, and Movement (Sandler et al., 2011). In their study of an emerging village sign language, Al-Sayyid Bedouin Sign Language (ABSL), Sandler et al. (2011) found that ABSL did not exhibit duality of patterning. When compared with established sign languages ABSL exhibits many more gesture types. For instance, ABSL has a much larger inventory of handshapes than occurs in established sign languages (see Sandler et al., 2011, Table 3). Many of the handshapes are, in typological phonological terms, highly marked, and would not be expected to occur in doubly articulated sign language phonologies. Sandler et al. (2011) suggests that the phonology-semantics interface is an emergent property of a growing sociocultural tradition of signing rather than the unmediated reflection of an innate cognitive-linguistic architecture. The takeaway from the ABSL data is straightforward: There appears to be a general tendency in the historical development of sign language phonology towards a progressive reduction of formal resources employed in signing concurrent with the development of duality of patterning (cf. Frishberg, 1975).

Writing systems also offer a fertile evidentiary source for conceptualizing what precedes the duality of patterning in the development of linguistic systems. In all three of the cases where writing is thought to have been independently invented—Mesoamerica, China, and Mesopotamia—writing systems began with primarily logographic orthographies (Boltz, 1994; Schmandt-Besserat, 1996). Syllabaries and alphabets, the writing systems which most closely parallel duality of patterning in speech, are always derived forms. Here again, change from a monoplanar system to a dually patterned one involves a massive reduction in the number of minimal segments of form employed.

In both emergent sign languages and emergent writing traditions, the formal organization of the initial state of linguistic varieties does not exhibit duality of patterning. Rather, distinct morphemes are represented by means of a wide variety of unique and perceptually distinct forms, whether manual or graphic. Analogs from signed and written linguistic varieties suggest that monoplanar speech was a historical precursor to duality of patterning in spoken natural language. Furthermore, in both the manual and graphic systems lacking duality of patterning, a much wider range of minimal units of representational form, whether handshapes or glyphs, are employed than in corresponding dually patterned systems. In terms of the evolution of speech, this would imply strong selection for complex articulations as components of very large inventories of monoplanar spoken languages. In particular, as we will now see, given the way in which velaric ingressive and pulmonic egressive airstream mechanisms combine to generate large numbers of complex segments, the addition of click consonants would offer the possibility of greatly expanding the number of semantic sense distinctions that could be made in monoplanar spoken language.

4. Multiplier Effect of Velaric Ingressives for Phonemic Inventory Size

The phoneme inventories of Khoesan languages are exceptionally large, hypertrophied in particular by their consonant inventories. Brugman (2009:28-29) cites the following numbers for phoneme inventories of select Khoesan languages: Nǀuu (Tuu) has 86 segments (73 of which are consonants); ǀGui (Khoe) has 99 segments (89 of which are consonants); Juǀ'hoansi (Ju) has 123 (89 of which are consonants). Maddieson (2005:10) finds a mean consonant inventory of 22.7 for a sample of 562 languages, with 6 for Rotokas (Papuan) and 122 for !Xóõ (Tuu) as the smallest and largest, respectively. Khoesan phonemic inventories are far and away the largest in the world, and this is due to their possessing click consonants.

Clicks are produced by means of a distinct airstream mechanisms than the one underlying the production of the vast majority of the most widespread phonetic articulations used in the world's languages. Phonation is typically achieved by modifying the outward flow of air. Phonemes produced in this manner are pulmonic egressives. Click consonants, however, are produced by lingual (or labial) suction. By making a closure with either the front of the

tongue or the lips, on the one hand, and the dorsum of the tongue, on the other, a low-pressure pocket of air can be created. When the front closure is released, high-pressure air from outside the mouth rushes in, creating the click sound. Phonemes produced in this manner are velaric (or lingual) ingressives.

Importantly, the addition of this other type of phonation does not just have an additive effect on phonemic inventory size—it has a multiplier effect. Because they rely on different airstream mechanisms, velaric ingressive consonants can co-occur synchronously with a range of pulmonic egressive phonations including voicing, nasalization, aspiration, glottalization, uvularization, and affrication (Clements, 2000:151). Take, for instance, the case of Juǀʼhoansi. There are four velaric ingressive consonant types in Juǀʼhoansi: dental, lateral, alveolar and palatal (ǀ, ǁ, !, and ǂ, respectively). Each of these types, however, can combine with 11 types of egressive phonation or 'efflux'. In !Xóõ, 5 click types (labial, dental, alveolar, lateral, and palatal) combine with 16 effluxes to make more than 80 different click types (Clements, 2000:151). It is the presence of clicks which makes these phoneme inventories so large. "Cross-linguistically no other sound class is subject to such an extensive series formation—inter alia by combining with a second, i.e., the pulmonic, air-stream mechanism" (Güldemann & Stoneking, 2008:105). Because of their multiplier effect on phoneme inventory size, in monoplanar spoken languages there would have been strong motivation for the elaboration of velaric ingressives.

5. Conclusion

Hockett, Martinet (1960:15), and Hjelmslev (1961:46) all argued that duality of patterning is an adjustment not strictly necessary for language, as such, but for the creation of an indefinitely large and open-ended lexicon. In his foundational piece on the evolution of speech, Charles Hockett (1960:96) wrote that "[t]here is excellent reason to believe that duality of patterning was the last [design feature of speech] to be developed, because one can find little if any reason why a communicative system should have this property unless it is highly complicated." Hockett's historical sequencing accords with evidence that adaptations necessary to produce complex speech patterns came relatively late in human evolution (MacLarnon & Hewitt, 1999; Lieberman, 2007). If duality of patterning relies upon a specifically evolved genetic endowment, we would expect selection pressure to be the greatest in the transition to speech; a much larger variety of perceptibly distinct segments can be produced in the manual-visual modality than in the oral-aural one; manual-visual codes lacking duality of patterning, like the Al-Sayyid Bedouin Sign Language studied by Sandler et al. (2011), can have large lexicons notwithstanding the limitations of their phonological organization.

A linguistic stage lacking duality of patterning would have represented a significant evolutionary bottleneck in the development of speech. Comparative evidence from the diachronic development of written and signed linguistic

varieties suggests that dually patterned phonological systems would have been preceded by a monoplanar stage in which phonological segments were mapped directly onto semantic functions. During a phase of monoplanar spoken language, any expansion in the number of perceptibly distinct sound segments produced would enable a comparable expansion in the lexical or morphemic inventory of the language. The use of a different airstream mechanism would have been invaluable here; velaric ingressives form highly productive series in conjunction with pulmonic egressives, much more productive than either implosives or ejectives. Indeed, in the core click languages, languages with the largest phoneme inventories in the world, click consonants routinely make up more than half of the inventory segments. Velaric ingressive, thus, would have been invaluable in monoplanar spoken languages.

The evolutionary development of duality of patterning would have enabled a newfound flexibility for phonological inventories hitherto operating at capacity. Nevertheless, the sound systems of those first dually patterned spoken languages would have converted segments sourced from extant monoplanar languages in fashioning their phoneme inventories. Given the importance of velaric ingressives for generating large inventories, clicks would likely have been derived segments in the first languages possessing duality of patterning. Velaric ingressives would thus have been an areal linguistic feature where dually patterned languages first emerged. The contemporary distribution of phonemic clicks offers supporting evidence for the deduction; velaric ingressives have a unique typological profile, being found in a linguistic area— southern and eastern Africa—where modern humans are thought to have originated (Henn et al., 2011).

This hypothesis, if true, would explain why phoneme inventories, but not other features of language, evince an (albeit noisy) "out-of-Africa" signal. Large phoneme inventories reflect, on this account, the influence of a historically prior, monoplanar phonological system. Large phonemic inventories are prevalent in southern African languages not because these have a more direct genealogical relationship with such monoplanar protolanguages, but because of sustained linguistic contact between language groups employing large and articulatorily complex phonological inventories in this area. The highly marked velaric ingressive airstream mechanism is best conceptualized as an areal feature kept alive as much through horizontal, as by vertical, transmission, and too marked, when compared to the unmarked pulmonic egressive airstream mechanism, to be independently developed outside of a sub-Saharan linguistic contact zone.

References

Atkinson, Q. (2011). Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa. *Science, 332,* 346-349.

de Boer, B., Sandler, W. & Kirby, S. (2012). New perspectives on duality of patterning: Introduction to the special issue. *Language & Cognition, 4(4),* 251-259.

Boltz, W. (1994). *The Origin and Early Development of the Chinese Writing System*. New Haven: American Oriental Society.

Bowern, C. (2011). Out of Africa? The logic of phoneme inventories and founder effects. *Linguistic Typology, 15(2),* 207-216.

Brugman, J. (2009). *Segments, Tones and Distribution in Khoekhoe Prosody*. Ph.D. Thesis. Cornell University.

Campbell, L. (2004). *Historical Linguistics*. Cambridge, MA: MIT Press.

Clements, G.N. (2000). Phonology. In B. Heine and D. Nurse (Eds.), *African Languages: An Introduction* (pp. 123-160). Camb.: Cambridge Univ. Press.

Creanza, N. et al. (2015). A comparison of worldwide phonemic and genetic variation in human populations. *PNAS, 112(5),* 1265-1272.

Donohue, M. & Nichols, J. (2011). Does phoneme inventory size correlate with population size? *Linguistic Typology, 15,* 161-170.

Frishberg, N. (1975). Arbitrariness and iconicity: Historical change in American Sign Language. *Language, 51(3),* 696-719.

Güldemann, T. & Stoneking M. (2008). A historical appraisal of clicks: A linguistic and genetic population perspective. *Annual Review of Anthropology, 37,* 93-109.

Hay, J. & Bauer L. (2007). Short report: Phoneme inventory size and population size. *Language, 83,* 388-400.

Henn, B., et al. (2011). Hunter-gatherer genomic diversity suggests a southern African origin for modern humans. *PNAS, 108,* 5154-5162.

Hjelmslev, L. (1961). *Prolegomena to a Theory of Language*. F. Whitfield (Trans.). Madison: University of Wisconsin Press.

Hockett, C. (1960). The origin of speech. *Scientific American, 203,* 88-111.

Knight, A., et al. (2003). African Y chromosome and mtDNA divergence provides insight into the history of click languages. *Current Biology, 13,* 464-473.

Lieberman, P. (2007). The evolution of human speech: Its anatomical and neural bases. *Current Anthropology, 48(1)*, 39-66.

Lindblom, B. & Maddieson, I. (1988). Phonetic universals in consonant systems. In C. Li & L. Hyman (Eds.), *Language, Speech and Mind* (pp. 62-78). London: Routledge.

MacLarnon, A. & Hewitt, G. (1999). The evolution of human speech: The role of enhanced breathing control. *American Journal of Physical Anthropology, 109(2),* 341-363.

Maddieson, I. (2005). Consonant inventories. In M. Haspelmath et al. (Eds.), *The World Atlas of Language Structures* (pp. 10-13). Oxford: Oxford University Press.

Martinet, A. (1960). *Éléments de linguistique générale*. Paris: Armand Colin.

Newmeyer, F. (2002). Uniformitarian assumptions and language evolution research. In A. Wray (Ed.), *The Transition to Language* (pp. 359-375). Oxford: Oxford University Press.

Nichols, J. (2012). Monogenesis or polygenesis: A single ancestral language for all humanity? In M. Tallerman and K. Gibson (Eds.), *The Oxford Handbook of Language Evolution* (pp. 558-572). Oxford: Oxford University Press.

Perreault, C. & Mathew S. (2012). Dating the origin of language using phonemic diversity. *PLoS ONE, 7(4),* e35289.

Ringe, D. (2011). A pilot study for an investigation into Atkinson's hypothesis. *Linguistic Typology, 15(2),* 223-231.

Sandler, W. et al. (2011). The gradual emergence of phonological form in a new language. *Natural Language and Linguistic Theory, 29,* 502-543.

Schmandt-Besserat, D. (1996). *How Writing Came About*. Austin: University of Texas Press.

Tishkoff, S. et al. (2007). History of click-speaking populations of Africa inferred from mtDNA and Y chromosome genetic variation. *Molecular Biology and Evolution, 24(10),* 2180-2195.

Trudgill, P. (2011). Social structure and phoneme inventories. *Linguistic Typology, 15,* 155-160.