Language accepts letters from readers that briefly and succinctly respond to or comment upon either material published previously in the journal or issues deemed of importance to the field. The editor reserves the right to edit letters as needed. Brief replies from relevant parties are included as warranted.

Vowel targets without a hyperspace effect

March 17, 2004

To the Editor:

In a paper with broad implications published in this journal, Keith Johnson, Edward Flemming, and Richard Wright (hereafter, JFW; 'The hyperspace effect: Phonetic targets are hyperarticulated', *Language* 69.505–28, 1993) provided experimental evidence suggesting that listeners preferred vowels with more extreme values than those they produced. In their study, listeners were to pick a 'best exemplar' of a vowel from a systematic exploration of a vowel space created on a speech synthesizer. The chosen exemplars (averaged across male and female listeners) were more extreme in formant values than the average productions (by the males). JFW claimed this 'hyperspace' effect underlies all vowel production, even though the targets were more extreme than their speakers produced even when hyperarticulating. The analogy that they provide to motivate the hyperspace effect is that full vowels cannot be predicted from reduced vowels, so reduced vowels cannot be taken as the underlying form (p. 524). By assuming that careful speech, or hyperarticulation, is similarly unpredictable, JFW expected that vowel targets would be hyperarticulated.

In their perceptual study, this is what seemed to be happening: Preferred vowels were outside the articulatory space of the male talkers. We have elsewhere shown that the experimental data depends both on the acoustics of the synthetic stimuli used to measure the effect and on the dialect of the listeners (D. H. Whalen, Harriet S. Magen, Marianne Pouplier, A. Min Kang, and Khalil Iskarous, 'Vowel production and perception: Hyperarticulation without a hyperspace effect', *Language and Speech*, to appear, 2004). Here we summarize the results and further explore some of the theoretical reasons for finding that there is no need for a hyperspace effect.

First, for both studies, dialect affects production and thus shifts the vowel space. In JFW's data, the one non-California speaker shows that this is true. The production values for their speaker from New York are given separately; if the perception results are compared to the New York speaker's production space, only two of eleven vowels exhibit a hyperspace effect. The fronting of back vowels in the California dialect is responsible for much of the apparent hyperspace effect, as confirmed by the Rhode Island speakers in our study.

Second, in our study, as in JFW's, even when speakers hyperarticulated, their productions were less extreme than their perceptions. This would imply that hyperarticulations should lie on a line between the full vowel productions and the perceptual (hyperspace) target. In both sets of production results, the directions in which the hyperarticulations went in formant space were quite sensible. For our data, the direction from full to hyperarticulated vowel was away from centralization for five vowels, decentralized for either F1 or F2 for five vowels, and aberrant for /u/ (where the full and hyperarticulations were nearly the same). The perceptual targets, by contrast, were fairly randomly distributed relative to either full or hyperarticulated vowels. If hyperarticulation is supposed to be asymptotically closer to a 'hyperspace' target, then there is no evidence of that trend in either our data or that of JFW.

Third, the preferred perceptual targets of the female listeners were essentially identical to those of the male listeners, even though those targets were 'reduced' from the standpoint of the female vowel spaces (see Whalen et al.). What this indicates is that both sets of listeners were normalizing the vowel space of this synthetic 'talker'. By collapsing the perceptual results but not the production results, JFW failed to address the issue.

Fourth, the method used by JFW to test the perception seems to lead to more extreme judgments than listeners actually prefer. This 'method of adjustment' allows the listener to pick vowels from a grid until the best match for an English vowel (as specified by the keyword at the top of the computer screen) is found. Subjects are inclined to explore the whole space and to prefer the edges. We found this to be the case when the same stimuli were judged via a rating task rather than the method of adjustment. The rating task presented each stimulus separately for a judgment as to vowel identity and goodness of representing that vowel. These results were much closer to the

https://doi.org/10.1353/lan.2004.0156 Published online by Cambridge University Press

productions than the method of adjustment perceptions were. The appearance of a hyperspace effect, then, seems to have been caused by several factors.

These results show us that listeners are remarkably capable of adjusting to a speaker's vocal characteristics, even when that 'speaker' does not really exist. The formants used to create the stimuli for both JFW and Whalen et al. specified a vocal tract of a certain length, roughly that of an average male talker, with F0 values typical for a male talker. The perceptual result was a simulated talker whose formant values were somewhat more extreme than those of the average American talker, and listeners were quite consistent in identifying those vowels as such. To fully test the hyperspace proposal, it would be necessary to use a range of synthetic voices differing in the dimensions of the underlying vocal tract in the way that the speakers' vocal tracts vary. Even so, such a demonstration could be interpreted as showing only that vowel spaces are different, both in extent and location, and that listeners are sensitive to the information in the vowel signal that tells them what that speaker is like.

One implication of JFW's original proposal is that the true study of a language's vowel space depends not on acoustic measurements but on perceptual testing of preferred targets. However, JFW recommend only that hyperarticulated speech be elicited (pp. 525–26), despite asserting that it requires hyperspace targets to predict production. Their recommendation seems closer to the underlying state of affairs, and even the need for hyperarticulations is not well established. The notion that hyperarticulations cannot be predicted from full articulations just because full articulations cannot be predicted from reduced ones does not necessarily hold. At least for the point vowels, the physiological mechanisms that would make a narrower, and thus more extreme, constriction are fairly easy to predict given a full vowel. Languages tend to maximize both the clarity of a particular vowel and its distance from other vowels in the system.

JFW's claim that vowel spaces need a hyperspace description has far-reaching implications, including the elicitation issue just discussed. No production, even the hyperarticulated ones, would ever live up to the hyperspace ideal. It appears that the hyperspace effect is an artifact of the subjects and methodology used. Thus it seems that full vowels are still sufficient to describe hyperarticulated, full, and reduced vowels.

D. H. Whalen
[whalen@haskins.yale.edu]
Harriet S. Magen
[hmagen@ric.edu]
Marianne Pouplier
[marianne.pouplier@yale.edu]
A. Min Kang
[min.kang@yale.edu]
Khalil Iskarous
[iskarous@haskins.yale.edu]

Bichakjian revisited

May 4, 2004

To the Editor:

While I am reluctant to get involved in a debate whose fine points I have not studied at length, I think it important, in light of recent discussion on evolutionary forces in language (e.g. Frederick Newmeyer in *Language* 79.583–99 (2003) and Bernard Bichakjian in *Language* 80.2–3 (2004)), to make a few general points about an 'evolutionary approach' of language development and our understanding of language lineages:

1. If the 'evolutionary approach' means that linguistic changes 'on the whole have proceeded along a unidirectional course' and 'were probably driven by a Darwinian-like process whereby linguistic features . . . were continuously replaced with ever-more advantageous alternatives' (Bichakjian, p. 2), then we need to be told why the Indo-European languages exhibit widely diverse phonetological, morphological, and syntactic systems that stand in conflict with 'a unidirectional course'. While changes are advantageous in some neurological, auditory, or communicational sense (cf. my *Essays on time-based linguistic analysis*, Oxford University Press, 1996, p. 294), being advantageous does not necessarily trigger a change. How does one get around the problem that simplifying (however conceptualized) a sound segment, a morphological construct, or a syntactic pattern may simultaneously make an item in the context more complex and perhaps less advantageous? A change may violate the connatural implicational relationships between the item changed and some item in the context. Is being advantageous in sound systems based on 'ease of articulation' or on making an item fit more easily into a given environment? Complexity, which can be measured, can be more or less advantageous than something else. Epenthesis is evidently advantageous in one respect but more complex in a purely quantitative sense. Conversely, a deletion may leave incompatible items juxtaposed.

https://doi.org/10.1353/lan.2004.0156 Published online by Cambridge University Press

2. No general statements about change can hold water if they fail to distinguish natural (bioneurolinguistic) changes from abnatural (sociocommunicational) changes (ibid., 293), which disrupt a language system. Cf. spelling pronunciations like *often* with *t* vs. regular *soften* and *moisten* without *t*. It would be hard to conceptualize many abnatural kinds of changes as advantageous. Higher-level natural changes like assimilation, rule-generalization to new environments, etc. can create environments that are less natural; cf. the change of *s* to [z] at the beginning of a word in North German. If it is advantageous to demarcate a word boundary, why does South German not do that? More generally, what advantages accrue from regional variation?

3. It may forestall misconceptions about my position if I point out that I agree with Chomsky that language is in the brain and not 'out there' in social or geographical space. The bioneurolinguistic make-up of humans is responsible for what a language can and cannot be; and however linguistic change may be conceptualized, it has to comport with what the brain allows or even favors. Yet, it would be silly to contend that sociocommunicational contact with other systems—those of different regions, exponentially multiplied by class, age, style subsystems—do not impact on one another. In my opinion, the way language is embedded in the nervous system cannot be fully understood without attention to how the brain deals with the great number of variants that it handles so readily; see ibid., Ch. 4. The analysis of this feat REVEALS a great deal about the brain if natural patterns and changes reflect neural realities. While *st*, *sp*, and *sk* are more natural BEFORE a syllabic nucleus than *ts*, *ps*, and *ks* are, the mirror image prevails AFTER a tautosyllabic nucleus. If this pattern offers a clue to naturalness, it is the CHANGE of, say, *sk* to *ks* in *ask* that is decisive in showing whether *sk* or *ks* is more natural in the environment in question.

4. Isolated languages change little over long periods of time, as Icelandic and Polynesian prior to Capt. Cook suggest. Although isolation does not promote change, intimate language contact often yields a new language. Studies of a dozen daughter languages of English in Africa, the Caribbean, Melanesia, and Polynesia, stimulated by the early work of Derek Bickerton, Peter Mühlhäusler, and others in the last century, have shown that functor words drawn from either contact language are very freely used to calque the functors of the language whose syntactic system prevails. An example is offered by the set of words—*because of*, *on account of*, *in place of*, *during*, and others—which in 'Middle' English substitute for the Anglo-Saxon preposition *for*. Some of the items here are Anglo-Saxon; some are French. This counterexemplifies the shibboleth of many historical linguists that functors don't get calqued. The patterning of English compound prepositions is structurally French. Calquing DISGUISES the origin of the morphological or syntactic system for those who can only count lists and are heedless of system realities. Although the sound system typically comes from the underclass if the languages are spoken by different classes, the syntax usually begins with a syntax which, except for word order, seems to be near-universal for nascent creoles. If intimate contact prevails till the end of the creolization phase and the social situation does not get significantly altered, the structure of the syntactic and noninflectional morphology of a creole will come from the language of the overclass, though deviations (irregularities) will often be wiped out by an over-generalization of the pattern. That the noninflectional morphology of Romance -[*i*]*ty* is about as productive as Germanic *-ness* says a lot, as does the fact that causative *-ize* or *-ify* has replaced no longer productive Anglo-Saxon-derived *-en* (as in *moisten*). That English includes a significant number of French-derived syntactic patterns that would be unborrowable after a creolization phase says even more (see many examples in ibid., Ch. 10). Interestingly, *'nt* in *mustn't* functions like *ne* in *il ne faut pas* and unlike *nicht* in *muss nicht*. Obviously, words and sound systems are not the sole clue to a language's lineage. They are not even the best clue, especially when a positivistic analysis of sounds and formations amounts to no more than a list of unrelated observations—as in citable treatments.

5. Changes that create sounds that in isolation are more complex can be advantageous in a given context, presumably because of their greater neurolinguistic naturalness in producing or perceiving. In any event, the MIXED ORIGINS of languages is something that SHOWS why various Indo-European language families are so different in their different components—phonetological, morphological, and syntactic.

6. What many studies of language history and English grammar have failed to grasp is Willi Mayerthaler's theory of reversals in marked environments, where marked and unmarked are based on my formulation (ibid., 314) in terms of natural changes/patterns. Our understanding of naturalness is not based on abnatural changes or on normality—i.e. statistics; see ibid., Chs. 5–6, for numerous exam-

https://doi.org/10.1353/lan.2004.0156 Published online by Cambridge University Press

ples and details. Note that *you* is unmarked in indicatives; omitting *you* 'markers' (a term referring to outward form) unmarked imperatives; its presence markers marked imperatives. Mayerthaler's theory will account for why native speakers of my acquaintance who would never say 'Me gave it to she' nevertheless assert that it would be incorrect to say anything other than 'Sue and me/myself did it for she and he'. Like *me* in 'Me, I'm ready' and 'That's me', such natural uses are French-derived. Marilyn Monroe was made to look silly by pointing to her picture in a newspaper and saying 'That's I!'. When school grammars stopped proscribing 'That's me', they had no notion of how grammatically to justify their decision; it certainly cannot be justified in terms of Germanic syntax. English regularizes and augments the environments (beyond those in French) in which personal pronoun forms that include *me*, *us*, and *them* occur (examples in ibid., 200, 331f.) Obviously English, like French, doesn't have pronominal 'cases' in the sense of Gothic or Greek, etc. Chs. 5–6 just referred to make it clear why the semantic functions of *be gonna* and *will* as well as those of the passivizing auxiliaries *be* and *get* are reversed in irrealis environments—negation, interrogation, not-yet-real time, etc. For subordinate clauses, cf. Latin indirect discourse and notably subordinate clauses within that environment; equally significant data can be found in English subordinate clauses. If such reversals account for why school grammarians don't arrive at satisfactory accounts of these and many other phenomena, they also show how historical linguists have been misled by concentrating on everything EXCEPT syntactic systems—treating English (whose Anglo-Saxon vocabulary is said to be a mere 28% of the total) as derived *tout court* from Anglo-Saxon instead of from French. Ask native speakers of English how much longer it takes them to feel at home in the process of internalizing German and French syntax. (But don't ask the one out of a hundred that can acquire languages the way children do!) Then ask them which language's sounds are easier to master! I dare predict that the two questions will be answered differently.

Aloha, CH.-J. N. BAILEY
[OLP@orlapubs.com]

https://doi.org/10.1353/lan.2004.0156 Published online by Cambridge University Press