

Reply to Mahowald and Gibson and to Heggarty: No problems with short words, and no evidence provided

Mahowald and Gibson (1) suggest that the shorter word length of frequently used words, and not their stability, could mean that chance sound correspondences account for the pattern of results we report for cognate relationships among proto-words in our study of seven Eurasian language families (2). However, their -0.24 correlation between the phonological length of contemporary English words and our measure of cognate class size is not relevant to the question of chance sound correspondences among the proto-words from different language families. What must be demonstrated is that shorter proto-words were more likely to be judged cognate simply on the basis of their length.

In this regard, we find no empirical evidence that phonological word length influences judgments of cognacy among the proto-words in our sample. For example, we find no relationship between cognate class size and the phonological length of Indo-European proto-words, controlling for the words' frequency of use and their rates of lexical replacement in the Indo-European languages ($P = 0.56$). On the other hand, and

in agreement with results we report in our paper, rates of lexical replacement do predict cognate class size, after controlling for phonological word length and frequency of use ($P = 0.0003$). We found similar patterns for the six other language families we studied.

Heggarty (3) makes a number of assertions about the Languages of the World Etymological Database (LWED) and the impossibility of identifying deep ancestry, but does not provide empirical evidence to support his claims. This repeats the major obstacle to progress in the field—a reliance on assertions rather than clear statistical evidence. Providing such evidence was precisely the starting point of our paper: our previous research (4) empirically demonstrated the possibility of a small subset of words retaining traces of their ancestry long enough to connect languages or language families separated by more than 10,000 y. We then incorporated that work into a statistical model that makes predictions about the words most likely to show evidence of ancestry among the seven Eurasian language families, and these predictions were supported.

Mark Pagel^{a,b,1}, Quentin D. Atkinson^c, Andreea S. Calude^d, and Andrew Meade^a

^a*School of Biological Sciences, University of Reading, Reading, Berkshire RG6 6AS, United Kingdom;* ^b*Santa Fe Institute, Santa Fe, NM 87501;* ^c*School of Psychology, University of Auckland, Auckland 1142, New Zealand; and* ^d*Linguistics Programme, University of Waikato, Hamilton 3240, New Zealand*

1 Mahowald K, Gibson E (2013) Short, frequent words are more likely to appear genetically related by chance. *Proc Natl Acad Sci USA*, 10.1073/pnas.1308822110.

2 Pagel M, Atkinson QD, S Calude A, Meade A (2013) Ultraconserved words point to deep language ancestry across Eurasia. *Proc Natl Acad Sci USA* 110(21):8471–8476.

3 Heggarty P (2013) Ultraconserved words and Eurasiatic? The “faces in the fire” of language prehistory. *Proc Natl Acad Sci USA*, 10.1073/pnas.1309114110.

4 Pagel M, Atkinson QD, Meade A (2007) Frequency of word-use predicts rates of lexical evolution throughout Indo-European history. *Nature* 449(7163):717–720.

Author contributions: M.P., Q.D.A., A.S.C., and A.M. wrote the paper.

The authors declare no conflict of interest.

¹To whom correspondence should be addressed. E-mail: m.pagel@reading.ac.uk.