Commentary

On Consonants, Vowels, Chickens, and Eggs

Luca L. Bonatti,[1,2] Marcela Peña,[1,3] Marina Nespor,[4] and Jacques Mehler[1,5]

[1]International School for Advanced Studies, Trieste, Italy; [2]Université de Nantes, Nantes Atlantique Universités, Labécd (EA 3259), Nantes, France; [3]Laboratorio de Neurociencias, Escuela de Psicología, Pontificia Universidad Católica de Chile, Macul, Chile; [4]University of Ferrara, Ferrara, Italy; and [5]Laboratoire de Science Cognitive et Psycholinguistique, CNRS and EHESS, Paris, France

With the possible exception of Italian traffic regulations, any rule will generate a statistically detectable advantage for items instantiating the rule. Thus, although attempts to reduce structural phenomena (e.g., Peña, Bonatti, Nespor, & Mehler, 2002) to statistical computations (e.g., Seidenberg, MacDonald, & Saffran, 2002) have been unsuccessful so far (e.g., Endress & Bonatti, in press), it would be no surprise if one or another statistical measure would correlate with the structural phenomena under investigation. But would this mean the statistics caused the apparently rule-abiding behaviors, or are the statistics epiphenomena of underlying structures? Questions about chickens and eggs are always difficult to settle.

In language, vowels and consonants have widely differing properties. For example, vowels tend to harmonize and consonants to disharmonize; consonants are more numerous than vowels; and the brain localizations of consonants and vowels differ. We argued that these differences may indicate that vowels and consonants serve partially different roles: Consonants may contribute more to word identity and vowels to grammatical phenomena (Bonatti, Peña, Nespor, & Mehler, 2005). We predicted that participants exposed to streams of syllables would extract pseudowords relying more on patterns of consonants than on patterns of vowels, and our prediction was confirmed. Keidel, Jenison, Kluender, and Seidenberg (2007, this issue) observe that our experiments controlled for the *token* frequency of the consonant and vowel sequences we used, but did not explicitly control for the statistical distribution of these sequences throughout the language (*type* frequency). Indeed, *Lexique* attests that there are 1.46 times as many consonant as vowel sequence types in three-syllable words of alternating consonants (Cs) and vowels (Vs; i.e., CVCVCV). Thus, as Keidel et al. elegantly observe, consonant sequences carry more information than vowel sequences. Keidel et al. suggest that this statistic, rather than any language-specific structural difference, drove participants' responses in our experiment.

But is the causal direction from statistics to structure, or vice versa? If, as we argue, consonants are more prominent than vowels in word identification, then surely under some measure consonant sequences *have* to vary more than vowel sequences. Indeed, given that most natural languages have more consonants than vowels, languages with more vowel than consonant sequences are unlikely to exist for combinatorial reasons alone, and French already exhibits an exceptional consonant-vowel balance relative to most languages. As we wrote previously, this "clearly makes consonants relatively more informative than vowels" (Nespor, Peña, & Mehler, 2003, p. 205). So our thesis predicts what Keidel et al. observed.

Yet Keidel et al. suggest that we may have taken the effect for the cause—that vowels and consonants may be different because of a statistical accident, captured by mutual information. In addition to the linguistic reasons just outlined, empirical arguments suggest that this is not so. First, differences in mutual information may apply when consonants are compared with vowels, but not when consonants are compared with consonants or when vowels are compared with vowels. By testing words and part-words in our experiments, we pitted consonant sequences against consonant sequences (Experiment 1) and vowel sequences against vowel sequences (Experiment 2). Thus, differences in mutual information between vowels and consonants cannot explain why French participants, who surely have an overwhelming amount of experience with vowels, failed to compute differences in transitional probabilities between strings of vowels.

Second, the observation that consonant sequences carry more information than vowel sequences cannot explain why different languages, or slight variations in the experimental materials, generate different results. According to Newport and Aslin (2004), English participants compute transitional probabilities on both

Address correspondence to Luca L. Bonatti, SISSA/ISAS, 2-4 Via Beirut, 34014 Trieste, Italy, e-mail: lucabonatti@mac.com.

vowels and consonants, whereas we showed that French participants do not. However, the consonant sequences in CVCVCV strings carry even more information in English than in French.[1]

Finally, that consonants have a prominent role in word identification is only half of our hypothesis. The other half is that, because vowels carry grammatical information, they have a prominent role in rule extraction. In current work testing Italian speakers, we (Toro, Bonatti, Nespor, & Mehler, in press) have found that when a stream contains both "words" and a "rule," participants cannot find simple generalizations defined over consonant sequences, but do find generalizations defined over vowel sequences. Had participants responded as Keidel et al. suggest they would, they would have generalized over consonant sequences, which outnumber vowel sequences in CVCVCV strings to a far greater extent and carry more mutual information in Italian than in French.

Thus, although we admire demonstrations of powerful statistical abilities in humans, we remain convinced that it is the linguistic chicken that lays statistical eggs, and not the statistical eggs that hatch into linguistic chickens.

Acknowledgments—The research for this article was funded by a European Science Foundation Eurocores "The Origin of Man, Language and Languages" grant to Marina Nespor and Jacques Mehler, by Human Frontiers Science Project Grant RGP 68/2002, and by Friuli-Venezia Giulia Grant Legge Regionale 11/2003 to Luca Bonatti.

[1]In the CVCVCV forms coded phonologically in the English MRC Psycholinguistic Database (Wilson, 1988), consonant sequences outnumber vowel sequences at a ratio of 2.6. Surely other factors, including distributional ones, may distinguish English from French. However, our point is that mutual information alone does not suffice to explain the cross-linguistic differences.

REFERENCES

Bonatti, L.L., Peña, M., Nespor, M., & Mehler, J. (2005). Linguistic constraints on statistical computations: The role of consonants and vowels in continuous speech processing. *Psychological Science*, *16*, 451–459.

Endress, A., & Bonatti, L. (in press). Rapid learning of syllable classes from a perceptually continuous speech stream. *Cognition*.

Keidel, J.L., Jenison, R.L., Kluender, K.R., & Seidenberg, M.S. (2007). Does grammar constrain statistical learning? Commentary on Bonatti, Peña, Nespor, and Mehler (2005). *Psychological Science*, *18*, 922–923.

Nespor, M., Peña, M., & Mehler, J. (2003). On the different roles of vowels and consonants in speech processing and language acquisition. *Lingue e Linguaggio*, *2*, 201–229.

Newport, E.L., & Aslin, R.N. (2004). Learning at a distance I. Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127–162.

Peña, M., Bonatti, L.L., Nespor, M., & Mehler, J. (2002). Signal-driven computations in speech processing. *Science*, *298*, 604–607.

Seidenberg, M.S., MacDonald, M.C., & Saffran, J.R. (2002). Does grammar start where statistics stop? *Science*, *298*, 553–554.

Toro, J.M., Bonatti, L.L., Nespor, M., & Mehler, J. (in press). Finding words and rules in a speech stream: Functional differences between vowels and consonants. *Psychological Science*.

Wilson, M.D. (1988). The MRC psycholinguistic database: Machine readable dictionary, Version 2. *Behavior Research Methods, Instruments, & Computers*, *20*, 6–11.

(Received 1/18/07; Accepted 1/29/07;
Final materials received 3/12/07)