LETTERS TO *LANGUAGE*

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.

Grammar and usage: A variationist response

July 8, 2005

To the Editor:

The recent discussions of grammar and usage in the pages of *Language*—by Frederick Newmeyer (N) ('Grammar is grammar and usage is usage', 79.4.682–707, 2003; 'A reply to the critiques of "Grammar is grammar and usage is usage" ', 81.1.229–36, 2005), Brady Clark (C) ('On stochastic grammar', 81.1.207–17, 2005), Ritva Laury and Tsuyoshi Ono (LO) ('Data is data and model is model: You don't discard the data that doesn't fit your model!', 81.1.218–25, 2005), Charles F. Meyer and Hongyin Tao ('Response to Newmeyer's "Grammar is grammar and usage is usage" ', 81.1.226–28, 2005)—make little reference to the extensive work in variationist sociolinguistics addressing these issues. Worse, one overt reference (N 2005:229) clearly misrepresents a foundational work in this field—Henrietta I. Cedergren and David Sankoff's 'Variable rules: Performance as a statistical reflection of competence' (*Language* 50.2.333–55, 1974; CS).

Work in this tradition explicitly problematizes N's grammar/usage distinction, not by rejecting grammar, but by improving its adequacy. Variationist research seeks not merely to generate all and only the grammatical utterances, but also to adequately characterize their contextual appropriateness and likelihood. N apparently postulates that grammar must be limited to categorical principles; hence he rejects out of hand any stochastic or probabilistic elements in the grammar. Variationist work begins with the contrary postulates enunciated by Uriel Weinreich, William Labov, and Marvin Herzog ('Empirical foundations for a theory of language change', *Directions for historical linguistics: A symposium*, ed. by Winfred P. Lehmann and Yakov Malkiel, 95–188, Austin: University of Texas Press, 1968; WLH): inherent variability and orderly heterogeneity. Inherent variability, as C observes, depicts variation, rather than perfect invariant consistency, as the normal state of linguistic systems; orderly heterogeneity is WLH's term for the fact that much of this variability in language is highly structured, showing probabilistic regularities, if not categorical ones.

Based on these principles, variationist work offers integrative resolutions to the grammar/usage differences that N sees as irreconcilable. Limited as to space, I restrict my comments to three of these here: probabilistic grammar, the use of corpora, and the multiplicity of styles and genres.

The quantification of grammar follows from inherent variability. Speakers manifestly are sensitive to frequencies and manipulate them systematically, and a probabilistic grammar more adequately models such competence than a categorical one, while maintaining the capacity to incorporate categorical statements (these are the limiting cases of probability values at zero or one). N's position, that there are no weights in the grammar other than zero and one, is simply an assumption, one that is massively falsified by the data.

N's principal argument against quantification is the claim that statistical regularities are 'overwhelmingly epiphenomenal' (2003:696). To variationists, this is an extreme case of special pleading. In fact, the multivariate analytical techniques like VARBRUL used in this research specifically control for the uneven data distributions that N invokes, and the results of studies of probabilistic linguistic constraints on variation typically find explanations in terms of recognizable grammatical properties. Two examples will illustrate this point. First, coronal stop deletion in English is probabilistically favored by phonological similarity with the preceding segment (there is more deletion in tokens like *wrist* than *rift*, *slapped* than *slammed*), a probabilistic instantiation of the obligatory contour principle (OCP), which disfavors adjacent identical elements in phonological structures. The OCP is a well-established principle of phonology with many categorical effects. The relationship between probabilistic and categorical instantiations of the OCP is that some particular sequences are absolutely prohibited, while others are probabilistically avoided; but in the latter cases, the disfavoring is significant, not epiphenomenally derived from frequencies of occurrence of the segments and features involved. This is the issue that C refers to as the 'stochastic generalization'.

How does N propose to exclude these quantitative findings from the grammar? If the OCP

https://doi.org/10.1353/lan.2005.0132 Published online by Cambridge University Press

is a principle of grammar, necessary for adequate accounts of numerous phonological phenomena, and grammar is purely categorical, then the probabilistic instantiations of the OCP are what? Merely random, fortuitous findings? If so, why are they systematic? Or are they the product of some stochastic grammar of usage, which happens to include an OCP(probabilistic) which is independent of, although parallel in most respects to, the OCP(categorical) residing in a Newmeyer-esque grammar of competence? Keeping probabilistic information out of the grammar would require systematic duplication of all grammatical principles that have both categorical and probabilistic manifestations. Occam's razor argues against such an analysis.

Of course, N goes on to raise the hoary argument that perhaps syntax is different from phonology, but massive work on syntactic variation reveals the same kind of nonepiphenomenal 'stochastic generalizations' as the OCP case. As an example, consider variable subject-verb agreement in Brazilian Portuguese: agreement is disfavored by postposed subjects (e.g. *Sumiu os taxis* 'Disappeared.SG the taxis'; cf. plural *sumiram*). Some researchers explain this effect by postulating that subjects of unaccusative verbs are originally VP-internal, and must raise across the verb to trigger the feature checking that accomplishes agreement. Would N reject this explanation out of hand, simply because the facts are not categorical (i.e. postverbal subjects don't have 100% failure of agreement, only, say, 90%)?

In his reply to C's discussion of stochastic grammar, N seriously misunderstands the probabilistic hypotheses underlying variable rules (VR). He claims (2005:229) that CS explicitly state that 'numerical quantities associated with the features in the environment of a variable rule are not discrete probabilities ''in the head of the speaker'' ', which N apparently takes to mean that there is no actual quantification in a variationist grammar. The opposite is in fact true. What CS are emphasizing in the cited passage is that they do not claim that the SPECIFIC VALUES AND FORMULAS used in the model are necessarily homologous with the processes of the mental grammar; but they make it clear in the rest of the article that they do in fact postulate a quantified or probabilistic mental grammar. The sentence before the one N quotes states: 'The variable rules developed by Labov should, like other rules of generative grammar, be interpreted as part of individual competence' (335). In other words, variationist work DOES suppose noncategorical, quantified grammatical processes 'in the head of the speaker'. Contra N, C is therefore correct to locate work on stochastic grammar within the 'variable rules' approach. Few variationists would disagree with C's advocacy of 'models of mental grammars that incorporate probabilistic information' (207).

N's problem with the use of corpora implies ignorance of the corpus-based methodology of sociolinguistics. Yes, a corpus drawn from the *New York Times* does not offer a very close approximation to the linguistic input and personal grammars of Carl Pollard or Ivan Sag. But sociolinguistic corpora try to be faithful to socially realistic linguistic entities: speech communities. The relationship between the community and individual grammars is an explicit research focus in sociolinguistics, and the principal finding is that individuals are grammatically similar as a function of social proximity. N's thirty-year-old citation of Frank Anshen's 1975 'Varied objections to various variable rules' (*Analyzing variation in language: Papers from the second Colloquium on New Ways of Analyzing Variation*, ed. by Ralph W. Fasold and Roger W. Shuy, 1–10, Washington, DC: Georgetown University Press) to the contrary is unrepresentative. He might consult works such as my own 'Variation in the group and the individual: The case of final stop deletion' (*Locating language in time and space*, ed. by William Labov, 1–36, New York: Academic Press, 1980) and 'Competence, performance, and the generative grammar of variation' (*Variation, change and phonological theory*, ed. by Frans Hinskens, Roeland van Hout, and W. Leo Wetzels, 125–43, Amsterdam: John Benjamins, 1997). Hence, extrapolating within communities from group data to individual grammars IS justified. Granted, it's an approximation, but no more so than any research based on less than perfect sampling. Testing such approximations is what statistics are for. N's case reduces to an argument against poorly constructed corpora, with which sociolinguists would concur.

Finally, N's argument from multiplicity of genres also demonstrates unfamiliarity with variationist research. For the most part, stylistic variation is quantitatively simple, involving raising or lowering the selection frequency of socially sensitive variables without altering other grammatical constraints on variant selection; indeed, it is commonly assumed in VR analyses that the grammar is unchanged in stylistic variation. As for the bigger 'genre' differences that N cites, how does he propose to generate them if they cannot be subsumed within a single grammar?

https://doi.org/10.1353/lan.2005.0132 Published online by Cambridge University Press

In general, N's arguments on this point neglect the second fundamental observation of WLH: orderly heterogeneity. Many arguments like N's make this mistake: they assume that absence of categoricality means chaos, but in fact, linguistic diversity is well mannered and orderly, following observable principles of social and linguistic organization. It is not a case of *après grammaire, le déluge*, but rather, grammar and usage both exhibit structure and order, some of it probabilistic, some of it categorical.

In fairness, N is not alone in neglecting variationist work. LO state: 'Nearly twenty years ago, part of the field of linguistics started examining real utterances to see what they look like and what they tell us about the shape and the nature of grammar' (223). This implies that such work began after 1985. Apparently they were unaware of William Labov's 1969 'Contraction, deletion, and inherent variability of the English copula' (*Language* 45.4.716–62), CS 1974, or the fourteen years' worth of NWAV papers that predate the functionalist work they cite. I trust this letter contributes to expanding awareness of variationist work on the grammar/usage interface, as it were, whose foundations, after all, lie in articles published in *Language*.

GREGORY R. GUY
[gregory.guy@nyu.edu]

https://doi.org/10.1353/lan.2005.0132 Published online by Cambridge University Press