

Knowledge of grammar includes knowledge of syntactic probabilities*

SUSANNE GAHL

University of Illinois

SUSAN GARNSEY

University of Illinois

1. GRAMMAR AND LANGUAGE FORM.

1.1. WHAT GAHL AND GARNSEY (2004) SHOW. There is widespread consensus that one of the tasks of a grammar is to describe the form of language. What Gahl and Garnsey (2004; henceforth G&G) show is that the description of both /t,d/-deletion and word and pause durations in American English needs to make reference to statements of the form ‘such-and-such a syntactic configuration has such-and-such a probability of occurrence in this context’. Therefore, in order to account for the form of language, the grammar needs to include probabilities. In particular, we argue that verb subcategorization probabilities, or VERB BIASES, must be part of the grammar, since the match or mismatch between verb bias and syntactic context affects form.

In a critique of G&G, Newmeyer (2006; henceforth N) agrees that the description of /t,d/-deletion and of other aspects of pronunciation reflects probabilities. However, N argues that the relevant probabilities are ‘extra-grammatical’: ‘generalizations about usage lie outside of grammar per se’ (N, p. 2). We acknowledge that subcategorization probabilities are a function of USAGE—they are that by definition, in fact. But the fact that subcategorization probabilities are ultimately the result of language use is not at issue. What is at issue is whether speakers know subcategorization probabilities as part of their knowledge of grammar. To establish that this is the case, G&G provided evidence suggesting that subcategorization probabilities affect language form even in situations where only syntactic probabilities, not meaning biases, can guide language production.

In the discussion to follow, we first clarify a point about the experiment reported in G&G and briefly comment on N’s suggestions for future research on this issue. We then compare N’s meaning-based account of verb bias to the notion of syntactic probability in G&G and show that the latter correctly predicts facts about language production and comprehension that are unexplained under N’s account. Finally, we consider N’s suggestions for future research in more detail, arguing that the suggested research cannot shed any light on the issue at hand.

1.2. CONTROLLING FOR MEANING. Reiterating a point made in Newmeyer 2003, following Lavender 1978, N draws attention to the following fact: sentences with different syntactic structures differ in referential meaning, whereas phonological variants generally do not. As a result, any distributional differences among syntactic variants may result from meaning differences, rather than reflecting stochastic information in the grammar. For example, sentences in which a given verb, say, *confirm*, takes a direct object (*they confirmed the date ahead of time*) vs. a sentential complement (*they confirmed the date was correct*) differ in meaning, and hence observations about these sentences cannot be used as an argument for stochastic grammar. N argues on this basis that G&G’s argument is flawed. But it is N’s critique that is flawed here: G&G did not compare direct object (DO) sentences with sentential complement (SC) sen-

* We would like to thank Fritz Newmeyer for a very enjoyable exchange of ideas. The first author is grateful to Terry Regier for discussions of an early draft of this rejoinder.

tences. Doing so would indeed have made us vulnerable to the objection raised by Lavandera. Instead, we compared sentences with high-probability DOs to those with low-probability DOs, as well as high-probability SCs to low-probability SCs. The linguistic fact that allowed us to make this comparison was that particular syntactic contexts have different probabilities of occurrence following different verbs, depending on the verbs' bias. We compared the aggregate of high vs. low probability sentences WITHIN each type (DO or SC). We arrived at generalizations about pronunciation that pertained to sentences of a given type and a given probability, regardless of lexical content. In other words, verb biases allowed us to manipulate probabilities while keeping syntactic structure constant, thus isolating effects of probabilities from those of syntactic structure. We believe that N appreciates the fact that G&G studied phonological variation, not syntactic variation. However, an informal small-sample survey reveals that readers of N's remarks nevertheless walk away with the impression that G&G compared DO and SC complements to one another.

Oddly, the studies N suggests for future research are in fact vulnerable to the Lavandera objection: the hypothesized results N describes hinge on finding frequency effects that applied equally to all sentences or phrases with a given phrase structure (for example, all DOs), and which set that phrase structure apart from other structures (for example, SCs). But a comparison between DOs and SCs is exactly the kind of study that the objection applies to, since the two patterns may differ in meaning. We examine this issue in greater detail in §4 below.

Having clarified G&G's experimental design, we now turn to N's main objection. N argues that G&G fail to show that knowledge of language includes knowledge of probabilities associated with syntactic structures, because there are, according to N, systematic differences in meaning (more specifically, differences in real-world likelihood) between high- and low-probability sentences. We argue that speakers know syntactic probabilities, as part of their knowledge of grammar, AND real-world likelihoods.

2. VERB BIAS AND MEANING. N agrees that American English /t,d/-deletion and the duration of verbs and their complements, as reported in G&G, reflect systematic differences between high- and low-probability sentences. But, according to N, the probabilities in question are 'semantic': they are based on speakers' experience of using verbs in connection with 'things' and 'propositions'. In other words, N argues that the effects reported in G&G are the result of speakers' experience with the kinds of meanings typically expressed using particular verbs, not with the syntactic complements the verbs typically take.

N's objection is flawed, both conceptually and empirically. As N points out, most syntacticians agree that subcategorization depends on meaning and consider verb subcategorization frames to be part of the grammar. Yet, verb biases are just a continuous version of either-or subcategorization information. Since dependence on meaning does not render information unsuitable for inclusion in the grammar, there is no principled reason to include subcategorization information while excluding verb bias.

N claims that G&G did not appreciate the systematic connection between verb biases and meaning. In fact, G&G discuss the relation between verb bias and verb meaning at some length (p. 752) and refer the reader to studies exploring this relation further (Roland et al. 2000, Roland 2002, Hare et al. 2003). We may add here that verbs with apparently similar meanings do not always have similar biases. For example, while *propose* is DO-biased (.45 DO, .18 SC), *suggest* is SC-biased (.18 DO, .61 SC, according to the norming data in Garnsey et al. 1997). Similarly, while *assert* and *maintain* both

have a strong DO-bias (.64 and .74, respectively), *claim* has a low DO-bias (.06) and a high SC-bias (.69).

Under N's account, apparent effects of verb biases are really effects of meaning, not of probabilities over syntactic types. Under G&G's account, by contrast, speakers know probabilities of syntactic types. G&G's claim is that speakers' accumulated experience with verb usage comes to affect the processing of sentences with a given verb and a given syntactic structure, independently of the meanings being expressed.¹ These two proposals make testably different predictions. The available evidence strongly points to speakers' knowing syntactic probabilities, not just semantic ones. For example, evidence from eye movements during reading reflects these verbs' syntactic biases: verbs with similar syntactic biases pattern together, whereas verbs that are similar only in meaning do not (Garnsey et al. 1997). Under N's account, there is no explanation for this observation.

To gauge semantic probabilities as distinct from syntactic ones—that is, to guard against precisely the problem N is concerned about—G&G used plausibility ratings, a well-established if imperfect technique. Plausibility ratings reflect speakers' estimates of the relative likelihood of various scenarios. If raters believe that one scenario happens 'a thousand times more frequently' than another scenario, then they should—and generally do, contrary to N's footnote 2—rate the first scenario as more likely to happen than the second scenario. Studies evaluating this technique have concluded that such judgments are reliable and consistent (Connell & Keane 2004), and numerous psycholinguistic studies have used it (e.g. Trueswell et al. 1993, Trueswell et al. 1994, Garnsey et al. 1997, McKoon & MacFarland 2000, Gruber & Gibson 2004). N also objects to the use of verb bias as a measure of syntactic probability on the grounds that, as N speculates, bias-conforming sentences describe scenarios that are commonly the case, whereas bias-violating sentences describe scenarios that are less commonly the case. Ferretti and McRae (1999) used a plausibility-rating task that teases apart verb bias and frequency of scenarios: Ferretti and McRae asked participants to rate (on a scale from 1–9) 'how common is it for an X to do Y?'. For example, 'how common is it for someone to confirm a date?'. The results lend further support to the notion that bias-conforming sentences do not generally describe scenarios that hold more commonly than bias-violating ones.

In sum, plausibility and verb bias are not just conceptually independent, they also make separate, distinguishable contributions to language comprehension and production.

3. PIES AND FLUENCY. N suggests that the results reported in G&G fall under the generalization that 'the more often we do something, the faster we are able to do it'

¹ An anonymous referee raises the question of whether accumulated experience with verb bias reaches a plateau during the lifetime of a speaker, and if so, when. To our knowledge, this question has not been studied systematically. We believe that, while verb biases are probably fairly stable in adult speakers, the adult language processing system remains malleable, so that verb biases can in principle continue to change throughout a speaker's life. For example, both authors remember a time when we were unfamiliar with the 'statistical' use of the verb *regress*, as in *regress Y on X*, and would only use that verb intransitively. By now, the transitive use of *regress* is by far the most common in our own usage. We therefore believe that our processing systems have come to treat this verb as a transitive-bias verb. Research on phonological and syntactic processing supports the idea that the language processing system remains malleable in adults (Dell et al. 2000, Onishi et al. 2002, Kaschak & Glenberg 2004).

(p. 8). Under N's view, the results in G&G are the result of speakers' handling high-probability sentences faster than low-probability ones. This characterization of our results is false. N's generalization would lead one to expect higher speaking rates in high-probability sentences than in low-probability ones, but that is not what we found. N's analogy to his experiences with pies is misleading, since the observed effect was not one of fluency or of increased speaking rate in high-probability material. On pp. 765–66 of G&G, we discuss a number of additional arguments against fluency as the source of the observed effects. N is also mistaken in claiming that the subject NP in our sentences did not show an effect of usage probabilities (i.e. of the experimental manipulation in G&G). If the subject NPs had been 'immune' to the experimental manipulation, as N implies, then one would expect the duration of those phrases to remain constant across the various experimental conditions. Instead, there were significant differences. We did not have a specific prediction concerning the duration of those phrases. We suspect that the observed pattern partly reflects a compensatory relationship between the duration of subject NPs and verb phrases. In any case, the match between verb bias and syntactic structure clearly did affect the duration of those phrases, even if the nature of that effect remains to be explored further.

The research reported in G&G was a reaction to previous work on pronunciation variation that focused exclusively on string frequencies and item-to-item probabilities, such as the probability of a given segment, word, or phrase, given a preceding or following item. Such item-to-item probabilities concern sequences that could, in principle, be practiced and that could, therefore, show frequency effects as a result of increased fluency or processing speed. A principal motivation of G&G was the sense that the exclusive focus on item-to-item probabilities was doing probabilistic linguistics a disservice. Since there is broad consensus that knowledge of syntactic structure is part of grammar, it is important to ask whether syntactic probabilities affect the form of language: if no such effects could be found, that fact would cast doubt on the notion that grammar is probabilistic. Tracking down such effects was the contribution of G&G.

4. FUTURE RESEARCH. We applaud N's willingness to state what sorts of findings he would consider relevant to the debate. N describes two sets of hypothetical findings. For both sets, alternative explanations exist that draw on well-established assumptions about grammar, making the hypothesized patterns poor arguments in support of the more controversial notion of stochastic grammar. Moreover, even in the absence of alternative explanations, both sets of findings would be vulnerable to the objection that N mistakenly levels against the evidence offered in G&G. The first set of findings concerns direct objects (DOs) and sentential complements (SCs). According to N, if it could be shown that the rate of /t,d/-deletion differed in these two contexts, then that 'would be an impressive argument in support of stochastic grammar' (p. 11). Presumably, showing this would mean showing that the two structures differed in frequency, and that /t,d/-deletion was more common in the more frequent one. In fact, DOs are vastly more frequent than SCs;² and there are reasons to believe that the rate of verb-final /t,d/-deletion is higher in DO-contexts than in SC-contexts, given their prosodic characteristics (cf. Wightman et al. 1992, Shattuck-Hufnagel & Turk 1996; and cf. Cho 2004). G&G did not observe such an effect, but if we had, we would have been ill-

² About 4.5 times more frequent in Gahl et al. 2004, a database that includes a disproportionate number of SC-taking verbs.

advised to consider it conclusive evidence in support of stochastic grammar, given that prosody offers a far less controversial explanation. What factors shape prosody is of course a deeply controversial issue, but seeking to explain pronunciation variation in terms of prosody is an uncontroversial move.

Even in the absence of alternative explanations, one could not rule out real-world likelihood as potential confounding factors in the research sketched by N: analogously to N's reasoning, one could speculate that DOs occur more frequently than SCs because people remark more often on experiences involving 'transitivity' (someone or something affecting 'things') than those involving 'propositions'. Therefore, overall differences in DOs vs. SCs—or any other pair of syntactic patterns that differed in frequency—could be due to real-world likelihoods of the kind N discusses.

The same reasoning applies to the second scenario N describes, concerning the various patterns of the form NP-BE-_[AP A-to-V]. The frequency of any syntactic pattern, however 'resistant to semantic characterizability' (N, p. 12) is ultimately the result of how often people deploy the syntactic pattern in question. That fact makes it possible to claim, for any syntactic pattern, that its frequency is the result of real-world likelihoods.³ A way to avoid that objection is to examine sets of sentences with identical syntactic structures but different usage probabilities—this is the approach taken in G&G.

5. CONCLUSION. What justifies the inclusion of certain facts in the grammar, to the exclusion of others? While we do not endorse the presumption of modularity that underlies this question, we stand by the claim that the evidence presented in G&G warrants including syntactic probabilities in the grammar.

REFERENCES

- CHO, TAEHONG. 2004. Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics* 32.2.141–76.
- CONNELL, LOUISE, and MARK T. KEANE. 2004. What plausibly affects plausibility? Concept coherence and distributional word coherence as factors influencing plausibility judgments. *Memory & Cognition* 32.2.185–97.
- DELL, GARY S.; KRISTOPHER D. REED; DAVID R. ADAMS; and ANTJE S. MEYER. 2000. Speech errors, phonotactic constraints, and implicit learning: A study of the role of experience in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 26.6.1355–67.
- FERRETTI, TODD R., and KEN MCRAE. 1999. Modeling the role of plausibility and verb-bias in the direct-object/sentence complement ambiguity. *Proceedings of the 21st Annual Conference of the Cognitive Science Society*, ed. by M. Hahn and S. C. Stoness, 161–66. Hillsdale, NJ: Erlbaum.
- GAHL, SUSANNE, and SUSAN M. GARNSEY. 2004. Knowledge of grammar, knowledge of usage: Syntactic probabilities affect pronunciation variation. *Language* 80.748–75.

³To make the logic of the research proposed by N more explicit: N's proposal involves showing the pronunciation of these patterns to be similar across all semantically different subtypes, and then showing that the pronunciation was 'a function of the probability of the use of the adjective-complement structure vis-à-vis other structures in the language which are equally resistant to semantic characterizability' (N, p. 12). The first step—evidence of similar pronunciation across all such types—would show that semantic role differences do not necessarily affect pronunciation—just as expected, given what is currently known about prosody. The second step would invite the objection that N raises: suppose we found a syntactic pattern XYZ that was at least twice as frequent as all NP-BE-_[AP A-to-V] patterns, and suppose further that sentences with that pattern had twice the /t,d/-deletion rate of NP-BE-_[AP A-to-V] sentences—nothing would stand in the way of concluding that the differences in pronunciation were the result of speakers' having encountered XYZ sentences more often than NP-BE-_[AP A-to-V] sentences, as a result of real-world likelihood of the aggregate of all NP-BE-_[AP A-to-V] sentences.

- GAHL, SUSANNE; DANIEL JURAFSKY; and DOUGLAS ROLAND. 2004. Verb subcategorization frequencies: American English corpus data, methodological studies, and cross-corpus comparisons. *Behavior Research Methods, Instruments, & Computers* 36.3.432–43.
- GARNSEY, SUSAN M.; NEAL J. PEARLMUTTER; ELIZABETH MYERS; and MELANIE A. LOTOCKY. 1997. The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory & Language* 37.1.58–93.
- GRUBER, JEFF, and EDWARD GIBSON. 2004. Measuring linguistic complexity independent of plausibility. *Language* 80.583–90.
- HARE, MARY L.; KEN MCRÆ; and JEFFREY L. ELMAN. 2003. Sense and structure: Meaning as a determinant of verb subcategorization preferences. *Journal of Memory & Language* 48.281–303.
- KASCHAK, MICHAEL P., and ARTHUR M. GLENBERG. 2004. This construction needs learned. *Journal of Experimental Psychology: General* 133.3.450–67.
- LAVANDERA, BEATRIZ R. 1978. Where does the sociolinguistic variable stop? *Language in Society* 7.171–82.
- MCKOON, GAIL, and TALKE MACFARLAND. 2000. Externally and internally caused change of state verbs. *Language* 76.833–58.
- NEWMAYER, FREDERICK J. 2003. Grammar is grammar and usage is usage. *Language* 79.682–707.
- NEWMAYER, FREDERICK J. 2006. On Gahl and Garnsey on grammar and usage. *Language* 82.XX–XX.
- ONISHI, KRIS; KYLE CHAMBERS; and CYNTHIA FISHER. 2002. Learning phonotactic constraints from brief auditory experience. *Cognition* 83.B13–B23.
- ROLAND, DOUGLAS WILLIAM. 2002. *Verb sense and verb subcategorization probabilities*. Boulder: University of Colorado, Boulder dissertation.
- ROLAND, DOUGLAS; DANIEL JURAFSKY; LISE MENN; SUSANNE GAHL; ELIZABETH ELDER; and CHRIS RIDDOCH. 2000. Verb subcategorization frequency differences between business-news and balanced corpora: The role of verb sense. Paper presented at the Workshop on Comparing Corpora, Hong Kong, October 2000.
- SHATTUCK-HUFNAGEL, STEFANIE, and ALICE E. TURK. 1996. A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research* 25.2.193–247.
- TRUESWELL, JOHN C.; MICHAEL K. TANENHAUS; and SUSAN M. GARNSEY. 1994. Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory & Language* 33.3.285–318.
- TRUESWELL, JOHN C.; MICHAEL K. TANENHAUS; and CHRISTOPHER KELLO. 1993. Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, & Cognition* 19.3.528–53.
- WIGHTMAN, COLIN W.; STEFANIE SHATTUCK-HUFNAGEL; MARI OSTENDORF; and PATTI J. PRICE. 1992. Segmental durations in the vicinity of prosodic phrase boundaries. *Journal of the Acoustical Society of America* 91.3.1707–17.

University of Chicago
 Department of Psychology
 5848 S. University Ave.
 Chicago, IL 60637
 [gahl@uchicago.edu]
 [sgarnsey@psych.uiuc.edu]

[Received 20 October 2005;
 accepted 4 December 2005]