J. Linguistics **43** (2007), 695–700. © 2007 Cambridge University Press
doi:10.1017/S0022226707004811 Printed in the United Kingdom

NOTES AND DISCUSSION

Variability and modularity: A response to Hudson[1]

DAVID ADGER

Queen Mary, University of London

(Received 15 May 2007)

1. Introduction

Hudson begins his response to Adger (2006) (Combinatorial variability) with the following question: 'are the findings of variationist sociolinguistics compatible with standard versions of Minimalism' (p. 683) and ends with the statement that a rapprochement between variationist sociolinguistics and 'any theory which assumes that language is a "purely linguistic" module free of all social and usage information' is 'inherently impossible' (p. 692). In Adger (2006), I provided an analysis of the variable patterns of use of *was* and *were* in Buckie English which did indeed assume a 'purely linguistic' module, free of social and usage information. I assumed that speakers could have such information about lexical items, but that it was not accessible to the processes that put words together into phrases, or that establish agreement relations or dependency relations between words and between phrases. In that sense, I defended a model of Socio(linguistics)-free Syntax and Use-free Syntax, with the same kind of general architectural claims as models of Phonology-free Syntax (see, e.g. Zwicky & Pullum 1983).

The first point to address, then, is whether the findings of variationist sociolinguistics are indeed compatible with such a model. I think the answer is that they straightforwardly are, and I'll outline the general model I've been proposing immediately below, showing how this is trivially the case. The second question is then whether the analysis I put forward for the Buckie case supports the general claim. Note that even if it didn't, the architectural points I make still stand: the model is in principle compatible with variability. However, I'll argue that Hudson's criticisms of Adger (2006) are not convincing.

[1] This paper was written immediately after the conference Formal Approaches to Variation in Syntax at York, UK (May 2007), and the discussions I had there have helped my thinking on these issues enormously. Thanks especially to Joan Bresnan, Devyani Sharma and George Tsoulas, and to Jenny Cheshire, Dick Hudson and Jennifer Smith for comments on a written draft.

https://doi.org/10.1017/S0022226707004811 Published online by Cambridge University Press

2. A MODEL OF VARIABILITY

I argued for a model of grammar (**G**) which produces a POOL OF VARIANTS (PoV), where each variant is a distinct feature complex, with the same semantic interpretation, and with potentially different phonological forms.

(1) $\mathbf{G} \rightarrow \{v_1, \ldots v_i \ldots v_n\}$ (=PoV)

I also assumed a distinction between knowledge of language and use of language (Chomsky 1965), explicitly stating that I took the grammar **G** to be embedded in a performance model (see, especially, footnote 10, p. 527). One can conceive of the systems of use **U** as a choice function on the pool of variants, given a context of utterance **C**:

(2) $\mathbf{U}(\text{PoV},\text{C}) = v_i \in \text{PoV}$

The function **U** is extremely complex, and is sensitive to all sorts of properties of the elements of PoV: their phonology, their sociolinguistic connotations, whether they have been encountered recently, their frequency of occurrence in the life of the language user who is speaking, whether the language user likes that particular word, etc. It is also sensitive to many aspects of the context of utterance: the information structure of the discourse, pragmatic expectations about the interlocutor's knowledge, social expectations about appropriateness etc.

One further factor, which I made use of in Adger (2006), is that the phonological form of variants might not be evenly distributed across PoV, so that more than one v_i could have the same pronunciation. In this way, the structure of PoV itself has an effect on the phonological form of any particular utterance. I make this clear in my paper in the passage Hudson quotes:

> The grammar produces variants in a way that predicts particular probability distributions, but those probabilities can be perturbed at the point of use by factors such as ease of lexical access, recency effects, metalinguistic or social judgments on the form, etc. (p. 506)

The probability of any particular v_i being chosen in any speech event is a function of the factors which enter into the specification of **U** and **C**. The fact that some phonological form might be more common in a corpus than some other phonological form depends on both the factors specifying **U** and **C** and the structure of PoV itself. So we must distinguish between the probability of any variant, and the frequency of the particular phonological forms (note how this is subtly different from the classical Labovian notion of linguistic variable). The probabilities of the variants can be given as is standard in probability theory.

(3) $p(v_1) + \ldots + p(v_i) + \ldots + p(v_n) = 1$

https://doi.org/10.1017/S0022226707004811 Published online by Cambridge University Press

Each v_i here is a pair of syntactic features and phonological form. The probability of occurrence of any particular v_i can be empirically modeled in the way familiar from variationist sociolinguistics (e.g. Labov 1972). One could attach probabilities directly to the variants (in an objective probability model), although my hunch is that a Bayesian subjective probability approach is a more useful conceptual framework allowing dynamic updating of **U**.

The claim I made in Adger (2006) was that, in addition to (3), if there are more variants with one form (say *was*) than another (say *were*), then that will further impact on the frequency of *was* vs. *were* in a corpus. I then showed how this seems to be correct when one takes a large enough sample from a single speech community (see next section for further comment).

In this model **G** is Socio-free and Use-free. There is never any rule of grammar that makes reference to frequency of a variant, or to social status of a variant. Grammar is sensitive only to syntactic features and structures built up from them. **U**, on the other hand, is neither Socio-nor Use-free; however, **U** does not construct syntactic representations or constrain dependency relations between constituents. Moreover, this general model sees **U** as a dynamically changing function, responsive to the particularities of the utterance situation, and taking into account all of the multifarious factors that influence the particular choice of variant and that are the subject matter of much sociolinguistic research (hence the point about subjective probabilities immediately above).

One might not like this model, but there is no doubt that it contains a 'purely linguistic' module and that it is compatible with the fact of sociolinguistically conditioned variation. I think that this is a straightforward falsification of Hudson's basic claim.

A secondary issue is whether this is a plausible model, given our current knowledge of how language works and how it is situated in the general architecture of the mind. This is clearly an open question (unlike the compatibility claim I have just discussed), but one which I am happy to pursue via an elaboration of the model and by confronting the model with empirical and theoretical challenges.

Hudson proposes an alternative model in which

> syntactic patterns (among others) are learned inductively on the basis of experience, with a great deal of very specific information stored in memory about patterns such as subject-verb pairs (Goldberg 2006). This is exactly what the Buckie data show: that memorised items include *we was*, *you was*, *we were*, *you were*. Moreover, in this usage-based account, our memories of tokens may include their contextual specifics, such as who uttered them and when … (p. 693)

In Adger & Smith (2005) it is pointed out that the constraints on Buckie *was*/*were* variation, specifically the categorical impossibility of *was* with a

https://doi.org/10.1017/S0022226707004811 Published online by Cambridge University Press

third person plural pronominal subject (**they was*) is true for present tense *be* (**they is*) and also for main verbs (**they walks*). Stating this as memorisation of tokens loses the generalisation that third person plural pronouns do not allow variability in their agreement. This is something which rather needs to be stated as a grammatical fact. I can see how a usage-based approach might do this by extracting a generalisation about agreement and third person pronouns from a set of exemplars, but Hudson provides no argument that this is a BETTER way to approach the data than the analysis I proposed in Adger (2006). It is interesting that Smith, Durham & Fortune (2007) report that children acquiring this variable pattern acquire the categorical impossibility of third person pronouns with variable agreement as soon as they acquire the agreement system in general, and with no analogical leveling. I'd be interested to see how a usage-based account would capture this.[2]

3. The putative flaws

Hudson puts forward a case that the arguments presented in Adger (2006) were flawed in a number of ways. The first part of this case is that middle-aged women have a markedly lower rate of use of we was than the other groups. I suggested in personal communication to Hudson that this was possibly because of prescriptive knowledge. Hudson responds that 'the trouble with this argument is that it is no stronger than the arguments for excluding prescriptive knowledge from the grammar' (p. 687). He then says that it is always possible to exclude such knowledge from grammar, but that it is difficult to justify this move linguistically or psychologically. However, the onus is on Hudson here, since he is arguing that a Socio-free account is impossible. I completely agree with him when he writes 'performance is heavily influenced not only by the subject (*you* or *we*) but also by the speaker's age and sex in a complex interactive pattern' (p. 688) but this does not constitute an argument that age and sex are part of a speaker's grammatical knowledge, as we saw in the previous section.

The second part of the case is that the speakers, once categorised into age and sex groups, rarely show the precise pattern predicted by the mechanisms of Combinatorial Variability specified in Adger (2006). But a major point

[2] Jenny Cheshire points out to me that Buckie is possibly an unusual case in having the categorical/variable splits discussed in Adger & Smith (2005) and in being so little influenced by the standard. The model I defend here is compatible with grammars that display no such splits: G would play no interesting role; the PoV would just include phonologically different variants learned on the basis of experience, and **U** would be sensitive to how standard a particular variant is. The question is whether other models provide non-stipulative accounts of the difficult cases like Buckie, where the grammatical generalisations seem to require a fair amount of abstraction from the surface forms and where categorical/variable splits emerge in many places in the system.

https://doi.org/10.1017/S0022226707004811 Published online by Cambridge University Press

of Adger (2006) is that this pattern only emerges once enough data is put together. I explicitly discuss the fact that various factors will impact directly on the use of a particular variant in any speech event. My question was whether one could see a general pattern emerging when these factors were controlled for, and my suggestion was that such a pattern could be attributed to the structure of the pool of variants, and hence ultimately to the grammar. I argued that this was precisely what happened when we look at the patterns as a whole (see p. 527, where this is discussed). The lack of enough individual data was precisely the reason that the paper is presented as a plausibility argument, rather than a knock-down argument.

This point is probably worth emphasising here. The grammar **G** predicts *n* variants for a particular meaning with a uniform probability distribution. This uniform distribution of variants does not predict a uniform distribution of phonological forms, since the phonological forms are not themselves always evenly distributed over the variants. Moreover, in any particular speech event, the speaker's choice of variant will be given by **U**, which is sensitive to speaker-internal properties such as intention, processing and memory and to (ultimately internalised) properties of the utterance context, such as who the interlocutor is and what conversation has gone before. Across sociolinguistically categorised groupings we may (or may not) see emergent patterns of higher or lower frequencies for particular forms. Collating the data together is one way to empirically bring out the effect of the uneven distribution of phonological forms over variants predicted by the theory.

The third part of Hudson's case is that I assume that every lexical item has an equal chance of being used; that is, I assume a uniform probability distribution in the set of variants. He claims that common experience shows this is not so (for example, whether you tend to use *somebody* vs. *someone* depends on formality). But this argument is backwards. It is an empirical finding that the distribution of *someone* and *somebody* is non-uniform, and a departure from the null hypothesis. As is well known in Bayesian probability theory, it is crucial to assume a prior probability distribution and I simply assumed a uniform distribution, the null hypothesis; it would have been a curious move to assume anything else. That this assumption led to the correct predictions is itself an interesting finding.

The final point Hudson makes is a somewhat technical one. He suggests that the feature specification I adopt for singular second person pronouns forces there to be a semantic distinction between generic and personal *you*. However, this isn't actually the case, since this distinction can be, and I think is, negotiated in the semantic/pragmatic component, with the generic meaning being a pragmatic weakening of the personal one (i.e. it means something like 'any relevant person'). The idea that φ-features in general are presuppositional and may be manipulated by the pragmatics is, I think, well-established (Cooper 1983, Schlenker 2003).

https://doi.org/10.1017/S0022226707004811 Published online by Cambridge University Press

4. Conclusion

Hudson aims to show that 'inherent variability can only be accommodated in a theory of language structure which can also accommodate usage information about typical speakers and about frequencies'. I actually don't disagree, except for the word 'structure'. From my perspective, the mechanisms that build grammatical structure are blind to information about social facts or facts about usage: syntax is Socio-free and Use-free. However, a general theory of language of course needs to take use into account, and the model outlined in Combinatorial Variability and further clarified above provides an architecture which can do this.

REFERENCES

Adger, David. 2006. Combinatorial variability. *Journal of Linguistics* 42, 503–530.
Adger, David & Jennifer Smith. 2005. Variation and the Minimalist Program. In Leonie Cornips & Karen Corrigan (eds.), *Syntax and variation: Reconciling the biological and the social*, 149–178. Amsterdam: John Benjamins.
Chomsky, Noam. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
Cooper, Robin. 1983. *Quantification and syntactic theory*. Dordrecht: Reidel.
Goldberg, Adele E. 2006. *Constructions at work: The nature of generalization in language*. Oxford: Oxford University Press.
Labov, William. 1972. *Language in the inner city*. Philadelphia, PA: University of Pennsylvania Press.
Schlenker, Philippe. 2003. Indexicality, logophoricity, and plural pronouns. In Jacqueline Lecarme (ed.), *Research in Afroasiatic grammar II: Selected papers from the Fifth Conference on Afroasiatic Languages, Paris, 2000*, 409–428. Amsterdam: John Benjamins.
Smith, Jennifer, Mercedes Durham & Liane Fortune. 2007. Community, caregiver and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change* 19, 63–99.
Zwicky, Arnold M. & Geoffrey K. Pullum. 1983. Phonology in syntax: The Somali Optional Agreement Rule. *Natural Language & Linguistic Theory* 1, 385–402.

Author's address: Department of Linguistics, School of Modern Languages, Queen Mary, University of London, Mile End Road, London E1 6NS, U.K.
E-mail: d.j.adger@qmul.ac.uk

https://doi.org/10.1017/S0022226707004811 Published online by Cambridge University Press