

- ASC 2-119-19. ELANGOVAN, A. & K. F. SHIM. 1998. Dietary phosphorus requirement of juvenile tiger barb, *Barbus tetrazona* (Bleeker, 1855). *Aquarium sciences and conservation* 2(1):9-19.
- RAVA 1519-151-12. SAITS, V. A. J. & R. H. MEDEMA. 2000. Checking out the G₂/M transition. *Biochimica et biophysica acta / Gene structure and expression* 1519(1/2):1-12.
- RM 13-137-46. HÖRDT, W., V. RÖHMELD & G. WINKELMANN. 2000. Fusarinines and dimethylamino acid, mono- and dihydroxamate siderophores from *Penicillium chrysogenum*, improve iron utilization by strategy I and strategy II plants. *Biometals* 13(1):37-46.
- RF 103-117-22. КУРТЕС, В. & В. РЕВЕЛЛ. 2001. Measuring consumer quality judgments. *British food journal* 103(1):7-22.
- CLAN 13-117-35. Educational Affairs Committee, American Association of Clinical Anatomists. 2000. A clinical anatomy curriculum for the medical student of the 21st century: Developmental anatomy. *Clinical anatomy* 13(1):17-35.
- DRR 128-135-44. DENÉMES, D., C. J. DECHESNE, S. VENTEO, F. GAVEN & J. RAYMOND. 2001. Development of the rat efferent vestibular system on the ground and in microgravity. *Developmental brain research* 128(1):35-44.
- FR 44-117-18. PARRINO, L. A., SMERIERI & M. G. TERZANO. 2001. Combined influence of cyclic arousability and EEG synchrony on generalized interictal discharges within the sleep cycle. *Epilepsy research* 44(1):7-18.
- EX 2-115-23. COLES, S. G. & M. J. DIXON. 2000. Likelihood-based inference for extreme value models. *Extremes* 2(1):5-23.
- FR 31-117-20. YUN, S.-H., T. ARIE, I. KAMERKO, O. C. YÖDER & B. G. TURGEON. 2000. Molecular organization of mating type loci in heterothallic, homothallic, and asexual *Gibberella / Fusarium* species. *Fungal genetics and biology* 31(1):7-20.
- HM 28-113-26. SCHMITZ, A. & L. T. WASSERTHAL. 1999. Comparative morphology of the spiracles of the Papilionidae, Sphingidae, and Saturniidae (Insecta: Lepidoptera). *Journal of insect morphology and embryology* 28(1/2):13-26.
- IMG 51-1120-29. FITZSIMONS, M. M., H. CHEN & M. H. FOSTER. 2000. Diverse endogenous light chains contribute to basement membrane reactivity in nonautoimmune mice transgenic for an anti-laminin Ig heavy chain. *Immunogenetics* 51(1):20-29.
- ICN 436-114-16. MACKLER, J. M. & N. E. REIST. 2001. Mutations in the second C₂ domain of synaptotagmin disrupt synaptic transmission at *Drosophila* neuromuscular junctions. *Journal of comparative neurology* 436(1):4-16.
- JMR 179-111-12. Reuss, L. 2001. Isolated-polarized epithelial cells as an experimental system for cell physiology studies. *Journal of membrane biology* 179(1):1-12.
- JSR 132-116-18. YASUNAGA, T., Y. SUZUKI, R. OHKURA, K. SUTOH & T. WAKABAYASHI. 2000. ATP-induced transconformation of myosin revealed by determining three-dimensional positions of fluorophores from fluorescence energy transfer measurements. *Journal of structural biology* 132(1):6-18.

Lacus Forum XCVIII (2002): 51-58

LINGUISTICS AS AN EMPIRICAL SCIENCE: THE STATUS OF GRAMMATICALITY JUDGMENTS IN LINGUISTIC THEORY

PATRICK J. DUFFLEY
Laval University

'ANYONE WHO HAS TAUGHT an introductory syntax course has had the experience of presenting an "ungrammatical" example only to be told by some smart-aleck about an unsuspected interpretation on which the sentence is quite normal' (McCawley 1982:78). A recent discussion with one of my colleagues, in which I had the pleasure of playing the part of the smart-aleck, led me to some serious reflection about the question of grammaticality judgments and their role in linguistic methodology. The conclusions of these reflections will be presented in this paper. My colleague and I were examining sentences (1)a and (1)b below, the first of which was claimed to be acceptable and the second not:

- (1) a. What did John hurt himself fixing?
b. *What did John hurt Bill fixing?

Although I could sense that there was some sort of difference in ease of interpretation between these two sentences in favour of (1)a, I was not at all happy with the suggestion that (1)b was 'less acceptable' or 'less grammatical' than (1)a. So I asked my colleague in what sort of context someone would say (1)a, and the description ran more or less as follows: 'the speaker knows that John was fixing things around the house yesterday and that he hurt himself while fixing something, but he does not know what that thing was and would like to be informed thereof'. This led to the logical rejoinder that if the speaker knew that John was fixing things around the house with Bill yesterday and that he hurt Bill while fixing some object whose identity the speaker was ignorant of but wished to know, then (1)b is a perfectly acceptable English sentence for obtaining that information.

The sort of discussion referred to above, which is typical in linguistic circles, leads one to wonder just what people (linguists included) are doing when they make so-called grammaticality judgments. It also raises the more basic question of why some linguists use such judgments as their prime source of data. What are they hoping to prove? Can grammaticality judgments provide the sort of evidence that these linguists are looking for?

It is not at all evident why an account of speakers' competence in understanding and producing language should be based on behaviour in a situation where they are doing neither, but rather are being asked to report their intuitions about the

acceptability of certain sequences of words. The motivations for following such a procedure can be boiled down to four main reasons:

First, by eliciting judgments we can examine reactions to sentence types that might occur only very rarely in spontaneous speech or recorded corpora. This is a standard reason for performing experiments in social science—observational study does not always provide a high enough concentration of the phenomena we are most interested in. A second, related reason for using grammaticality judgments is to obtain a form of information that scarcely exists within normal language use at all—namely, negative information, in the form of strings that are not part of the language. The third reason for using judgments is that when one is merely observing speech it is difficult to distinguish reliably slips, unfinished utterances, and so forth, from grammatical production. A fourth and more controversial reason is to minimize the extent to which the communicative and representational functions of language obscure our insight into its mental nature. (Schütze 1996:2-3)

With respect to the last reason, Schütze admits that it presupposes a particular view of grammatical competence as cognitively separate from other facets of language knowledge and use, and hence its validity depends on one's theoretical stance on the issue. It will be my contention in this paper that all of the other reasons except the first are also theory-dependent and that on top of this grammaticality judgments are practically worthless as scientific evidence, even if one accepts the theoretical presuppositions of generative grammar.

To start with the last point, it must be realized that the term 'grammaticality judgment' itself is in fact a misnomer. What is actually meant would be better expressed by the term 'acceptability judgment' (cf. Schütze 1996:26). Given Chomsky's (1965:10-11) definition of grammaticality as belonging to the sphere of competence, that is, of the ideal speaker's knowledge of his language, it makes no sense to speak of 'grammaticality judgments', since grammaticality is not accessible to people's intuitions: all a native speaker can do is judge a string's acceptability. So what people are actually doing when reacting to a sequence of words presented to them by a linguist is judging whether it seems acceptable to them or not. Is there any relation, then, between acceptability and grammaticality which would allow inferences to be made about the latter based on the former?

According to generative theory there is indeed a relation between these two concepts, whose nature is described in the following well-known passages:

Acceptability is a concept that belongs to the study of performance, whereas grammaticality belongs to the study of competence... Grammaticality is only one of the many factors that interact to determine acceptability. Correspondingly, although one might propose various operational tests for acceptability, it is unlikely that a necessary and sufficient operational criterion might be invented

for the much more abstract and far more important notion of grammaticality. (Chomsky 1965:10-11)

...linguistics as a discipline is characterized by attention to certain kinds of evidence that are, for the moment, readily accessible and informative: largely, the judgments of native speakers. Each such judgment is, in fact, the result of an experiment, one that is poorly designed but rich in the evidence it provides. In practice, we tend to operate on the assumption, or pretense, that these informant judgments give us 'direct evidence' as to the structure of the language, but, of course, this is only a tentative and inexact working hypothesis... In general informant judgments do not reflect the structure of the language directly; judgments of acceptability, for example, may fail to provide direct evidence as to grammatical status because of the intrusion of numerous other factors. (Chomsky 1986:36)

These observations show that Chomsky himself is aware of the indirectness of the link between acceptability (performance) and grammaticality (competence), as indicated by the reference to the 'numerous other factors' that interact with the grammar to produce acceptability. Just how numerous and uncontrollable these factors are is shown by the studies of Birdsong 1989 and Schütze 1996. They comprise things like:

- (a) the instructions given to the subjects doing the questionnaire
- (b) the order of presentation of the sentences submitted for speaker reactions (the first sentences in a questionnaire tend to be judged much more severely than the others; cf. Greenbaum 1973, 1976)
- (c) the effect of the repetition of an unacceptable structure leading people to accept it
- (d) judgment strategies (one does not know whether the subjects are using the same criterion to decide on acceptability)
- (e) modality and register (a written questionnaire already represents a fairly formal context for most speakers; cf. Greenbaum 1977)
- (f) how much time is given to the informants to react
- (g) context (is it easy to imagine a possible context for the sentences?)
- (h) meaningfulness (can people make sense of the sentences?)
- (i) length and complexity of sentence judged
- (j) frequency of constructions (less frequent structures are often judged unacceptable)
- (k) lexical content of items
- (l) rhetorical structure (structural parallelism renders certain sequences acceptable which are not perceived so otherwise; cf. Langendoen 1972).

I would add to this long list an even more basic and important factor: the very fact of asking a speaker to make an acceptability judgment is asking him to do something completely unnatural. What people rather are accustomed to doing with their

language is not making grammaticality judgments, but simply using it to express themselves. Add to all this the fact that one is not supposed to know what the grammar looks like—it is what the linguist is trying to determine—and you get a bag so mixed that its contents are impossible to sort out.

Why, then, have recourse to data whose connection with the object of one's hypothesis is so tenuous? This question takes us back to the motivations given in defence of this procedure—rarity of significant data in spontaneous speech, the need to obtain negative data, the facilitation of distinguishing performance errors from grammatical production, the separating out of the communicative and representational functions of language from its mental structure. When scrutinized more closely, all of these reasons except the first turn out to be products of the theoretical stance adopted by generative grammarians. I have already alluded to Schütze's admission that this is the case for the last reason. Regarding the facilitation of the distinction between performance errors and grammatical production, it should be fairly obvious that this alleged justification merely begs the question by presupposing some concept of grammatical production. Moreover, as shown by the enumeration given above of the many possible performance errors which can occur in the making of grammaticality judgments, the use of such judgments does not facilitate the identification of potentially extralinguistic factors which have an impact on the data but merely adds further factors to the list. This has led Schütze (1996:179–180) to observe that: 'In fact, it might appear that grammaticality judgments are the *worst* way to get at linguistic competence, as compared to production and comprehension, because they involve the interaction of many more factors.'

Under Schütze's pen, this is a merely rhetorical objection. He goes on to give two reasons why this does not constitute grounds for abandoning grammaticality judgments as a source of data:

1. while more factors are involved in such judgments, they 'might be less mysterious than those connected to language use (how could we ever define the 'understanding of a sentence' or 'communicative intentions' and how could we draw conclusions about grammaticality from them?);
2. grammaticality judgments provide an alternative path to the grammar (they are subject to different influences than language use is and so facilitate the search for the common core that underlies both types of behaviour, i.e., the grammar).

Neither of these reasons stands the test however. Concerning the first motive, one of the crucial factors impacting on grammaticality judgments is necessarily whether the subject can understand the sentence or not (cf. Schütze 1996:162), a fact which would seem to make natural language comprehension no more 'mysterious' than such judgments. Moreover, if one compares comprehension and grammaticality judgments in Schütze's own model (p. 175), one notes that both are determined by the same four factors—input, knowledge (general, contextual, etc), competence and parsing

strategies—with four other factors being added for grammaticality judgments. This only makes sense—in order to judge a sentence's acceptability, one must first comprehend it—but the facts certainly do not support the suggestion that the process of judging a sentence for grammaticality is any less obscure than that of comprehending the sentence in a natural context—quite the contrary. As for the argument that grammaticality judgments provide an alternative path to the grammar, it is much sounder methodology to begin with the cases where it appears probable that the fewest factors are involved before attempting to come to grips with the more complex cases. As Birdsong (1989:72) puts it, 'the hypocrisy of rejecting linguistic performance data as too noisy to study, while embracing metalinguistic performance data as proper input to theory, should be apparent to any thoughtful linguist.'

Schütze's second reason—the need to obtain negative data—brings us even closer to our objective of understanding why linguists of the generative school have such regular recourse to grammaticality judgments. The very fact of needing to discriminate between certain sequences of words that are 'part of the language' and other which are not implies a certain view of both grammar and language which is peculiar to generative theory. The citation below provides a capsule summary of this view:

A major objective of linguistic research is to construct a grammar capable of generating all the grammatical sentences and no ungrammatical ones. This research involves identifying the rules that allow speakers to determine which sentences of their language are well-formed and which are not. (O'Grady & Dobrovolsky 1987:103)

Particularly revealing in this quotation is the close relation made between the project of constructing a generative grammar and the search for the rules that allow speakers to judge which sentences are well-formed and which are not. This suggests that a transfer has taken place from the role the grammar is claimed to have in the theory to the role of the subject in a grammaticality judgment: just as the grammar determines what is well-formed and what is not, so the speaker confronted with a string of words in a questionnaire decides what is structurally good and what is not. However this is definitely not what people do when they comprehend what others are trying to say in a normal speech situation (nor is it, as we have seen above, what they are doing when they make grammaticality judgments). Such a view of grammar makes it an algorithm for performing structural grammaticality choices rather than an instrument for carrying on communication.

Examples of this procedure abound; to give a typical case, one might refer to Givón's studies on causative verbs in two articles entitled 'Cause and Control: On the Semantics of Interpersonal Manipulation' (1975) and 'The Binding Hierarchy and the Typology of Complements' (1980). In his discussion of the verbs *cause*, *make* and *have*, Givón claims that these English verbs may be scaled according to two semantic properties which are universally attested: (a) intended ('controlled') vs. unintended ('uncontrolled') causation; (b) 'mediated vs. direct causation' (1980:335).

cause is a 'noncontrol causation verb', *make* a 'direct control causation verb' and *have* a 'mediated control causation verb'. When the data supporting these generalizations is confronted with actual usage however, the suspicion arises that the data was fabricated to support the universal semantic properties rather than the later being inferred from an observation of usage. For instance, the claim that *make* denotes deliberately intended causation while *cause* evokes accidental causation is based on the purported contrast in acceptability between (2)a and (2)b:

- (2) a. John accidentally/inadvertently caused Mary to drop her books.
b. *John accidentally/inadvertently made Mary drop her books.

Actual usage shows, however, that the distinction between *cause* and *make* has nothing to do with intentionality. On the one hand, *cause* can denote a deliberate action, as in (3):

- (3) If a person has thoughtlessly or deliberately caused us pain or hardship...
(Brown U. Corpus 808 0470)

On the other hand, *make* can evoke unintentional causation, as in:

- (4) Other women—they only made me love you more.
(O'Neill 1955 [vol. 1]:130)

The analysis of *have* as denoting mediated causation, which is intentional but requires the intervention of a third party, suffers from a similar lack of support from the empirical data. Givón adduces the purported contrast in acceptability between (5)a and (5)b/c:

- (5) a. I had her lose her temper by sending John to taunt her.
b. ? I caused her to lose her temper by sending John to taunt her.
c. ? I made her lose her temper by sending John to taunt her.

Actual usage in this case would seem, however, to be exactly the opposite of Givón's judgements: (5)a makes no sense at all, while (5)b and (5)c are quite normal. *Have* is used in English to evoke getting someone to do something by exercising one's authority or control over them through a request or command, as in (6) below:

- (6) The teacher had me recite my poem in front of the class.

This does imply intentionality on the part of the causer and compliance on the part of the causee, but there is no idea at all of mediation by a third party suggested by the meaning of the construction illustrated in the sentence above.

The conclusion we have been led to, therefore, is that grammaticality judgments are not a reliable source of empirical data. What speakers are doing when they perform such judgments is appraising the acceptability of the utterances they are being asked to evaluate. Even if one accepts the hypothesis of a separate grammatical module constituting one of the important factors which determine acceptability, the causal link between acceptability and grammaticality does not allow one to make inferences from one to the other. If natural language production (i.e., performance) is viewed as inadequate data for inferring conclusions about grammaticality, the data provided by grammaticality judgments must be considered even less trustworthy. As a type of metalinguistic behaviour, these judgments are themselves just another sort of performance, and as such they are subject to even more confounding factors than natural utterance production.

One might palliate some of the drawbacks of grammaticality judgment data by the design of the questionnaire used to elicit them. As shown by Gries (this volume), the inclusion of experimental controls such as randomized presentation of sentences, inclusion of fillers, and clear exemplification and explanation of the required judgment process, can elicit reactions which correspond fairly closely to corpus data. Moreover, if sentences were presented to informants with a context, that is, a description of the communicative situation, then, in this more natural setting, one should obtain more reliable judgments of acceptability.

The fact remains, nevertheless, that sound methodology would advise one to first study language in its natural setting before placing speakers in an artificial situation and asking them to do something entirely different from everyday language use. The very nature of a questionnaire suggests a testing of the informants' ability to conform to some norm of expected behaviour, and triggers the reaction 'what should one say in this situation?'. Even if one were to succeed in eliminating this conditioned reflex—something highly unlikely in the present author's opinion—there still remains a hypothetical element inherent in the nature of a questionnaire: informants are being asked to answer the query 'what would one say in this situation?'. Neither question corresponds necessarily to what the speaker *actually* says in a given situation. Thus, in any case, one is driven back to actual usage as the final test of the explanatory capacity of any theory. Isn't what people actually say what we linguists are supposed to be explaining in the first place?

REFERENCES

- BIRDSONG, DAVID. 1989. *Metalinguistic performance and interlinguistic competence*. Berlin: Springer-Verlag.
CHOMSKY, NOAM. 1965. *Aspects of the theory of syntax*. Cambridge: MIT Press.
———. 1986. *Knowledge of language: Its nature, origin, and use*. New York: Praeger.
GIVÓN, TALMY. 1975. Cause and control: On the semantics of interpersonal manipulation. In *Syntax and semantics*, vol. 4, 59–91. New York: Academic Press.

- _____. 1980. The binding hierarchy and the typology of complements. *Studies in language* 4:333-77.
- GREENBAUM, SIDNEY. 1973. Informant elicitation of data on syntactic variation. *Lingua* 31:201-12.
- _____. 1976. Contextual influence on acceptability judgments. *Linguistics* 18:75-11.
- _____. 1977. Judgments of syntactic acceptability and frequency. *Studia linguistica* 31(2):83-105.
- LANGENDOEN, D. TERENCE. 1972. The problem of grammaticality. *Peabody journal of education* 50(1):20-23.
- MCCAWLEY, JAMES D. 1982. How far can you trust a linguist? In *Language, mind and brain*, ed. by Thomas W. Simon & Robert J. Scholes, 75-88. Hillsdale NJ: Erlbaum.
- O'GRADY, WILLIAM & MICHAEL DOBROVOLSKY. 1987. *Contemporary linguistic analysis: An introduction*. Toronto: Copp Clark Pitman.
- O'NEILL, EUGENE. 1955. *The plays of Eugene O'Neill*. New York: Random House
- SCHITZE, CARSON T. 1996. *The empirical base of linguistics: Grammaticality judgments and linguistic methodology*. Chicago: University of Chicago Press.

MORPHOLOGICAL PATTERNS OF TERMS IN NEUROSCIENCE AND PARTICLE PHYSICS

CAROLYN G. HARTNETT
College of the Mainland

HOW HELPFUL AND ACCURATE IS THE LEXICON OF SCIENCE? Scientists and teachers of science since Lavoisier two hundred years ago have stated that knowing science means knowing the language for its concepts. Igor Melchuk (2001) has discussed the disarray of terms in European science in the last century. The British science journal *Nature* recently showed the concern of science with its lexicons by beginning a weekly column entitled 'Words', reporting the first use of the word *scientist* in 1834 (Danielson). Recently, it reported that the Sudbury Neutrino Observatory in Ontario combined its new measurements with some from the Super-Kamiokande detector in Japan to show that neutrino particles have unexpected mass and sometimes switch their identities. When their identities change, should their names change? An editorial called the finding 'not a crisis for existing models, but a route to deeper ones' (*Nature* 2001a). It welcomed the implications for the theory called the 'standard model', GUT ('Grand Unified Theory'), TOE ('Theory of Everything'), and SUSY (supersymmetry), as well as the yet-unseen Higgs boson¹.

If the lay public theorizes that science is difficult only because of its vocabulary, then the reasons for a difficult lexicon ought to be examined. What evidence of difficulty appears in the morphology of terms in science in English, the leading language of science now? How helpful is the morphology of the lexicons of science? How useful are the many glossaries? Since sciences have different vocabularies, I examined separately the lexicons of two different fields that are currently making great strides. I report comparative tabulations of the morphological patterns of terms in the fields of particle physics and neuroscience. I give examples and conclude with reasons for difficulties, based on the ways scientists must work. All the terms I discuss are standard ones used in publications, not vernacular for informal conversation in the lab.

Particle physics deals with extremely rare tiny particles, too small for instruments to detect, usually predicted only by logical gaps in paradigms and identified only by their effects. They can appear to be in two places at once. The Heisenberg Uncertainty Principle holds that not everything, such as location and velocity, can be known simultaneously. 'Spooky' is how Einstein described one of his own thought experiments². Astrophysicist John Gribbin is dissatisfied with the terms *particles* and *waves*, which he calls metaphors. 'We call those objects particles, for want of a better name. What they really are, we do not know' (1998b:51-52). The Nobel Prize-winning physicist Steven Weinberg also expresses concern about names for the materials he studied: