Advancing Linguistics Between the Extremes: Some thoughts on Geoffrey Sampson's *Grammar without Grammaticality*

W. Detmar Meurers
The Ohio State University
dm@ling.osu.edu

To appear in
Corpus Linguistics and Linguistic Theory 3(1), 2007.

The ready availability of large corpora has opened a range of interesting possibilities for linguistic research, and Geoffrey Sampson's article illustrates that a corpus-based perspective can also motivate revisiting the general direction and methodology of linguistic research. At the same time, the discussion in Sampson's article pushes two issues to rather extreme conclusions, which I think could be useful to revisit here: i. the issue of grammaticality and its role in linguistic research, and ii. the development of a scientific linguistic methodology, addressing the role of introspection and alternative ways of obtaining and evaluating data.

Turning to the first issue, Sampson argues against the grammatical/ungrammatical distinction underlying generative linguistics, proposing that one can only describe language usage, which he pictures as paths in open grassland—some paths are more frequently taken, but any path is in principle possible. While the linguistic preoccupation with grammaticality arguably deserves to be revisited, throwing out grammaticality altogether and replacing it with a description of whatever language happens to have been used seems like

throwing out the child with the bath water. What the advent of generative linguistics provided was a focus on language as a system that can be modeled in a way making testable predictions. It replaced the backward looking description of traditional linguistics with a process aimed at establishing generalizations over the observed patterns, which could be used to make predictions about previously unseen data. But as soon as one turns from descriptions to generalizations and predictions derived from those, one expresses statements about which sentences can be licensed and which cannot. The move that turns linguistic theories into falsifiable scientific proposals thus is inseparably tied to the question which sentences are licensed by a particular theory and which are not, i.e., the issue of grammaticality.

Interestingly, Sampson's own, pioneering work on the linguistic annotation of the SUSANNE corpus showcases that it is meaningful to establish generalizations over lexical classes and phrase structure, and that the rules used in the annotation of the corpus data support a parsing system making decisions about which strings can be licensed, i.e., are grammatical. Or in the words of Sampson (2000): A parsing system "automatically produces analyses (according to some parsing scheme) of input language examples. A parsing scheme defines the target which a parsing system hits (or misses). The SUSANNE Corpus represents part of the definition of a parsing scheme."

Grammaticality thus serves an important function in constraining the infinite set of possible sentences based on a finite representation generalized from observations, i.e., a linguistic theory or grammar. Modifying Sampson's metaphor of language use as paths in an unobstructed prairie, it thus seems that language use is more akin to the paths in a mountain region—there still is a wide range of possible paths, but the terrain rules out certain ones for general use; by looking at the topography, one can thus make predictions about where new paths could arise in the future (e.g., alongside a river as opposed to going up the north face of a mountain).

We see an immediate benefit of characterizing language use in this way when we see how well it matches the well-established characteristics of language change: "Language change is not a completely random, unprincipled deviation from a state of pristine perfection, but proceeds in large measure in a remarkably regular and systematic fashion" (Hock 1991, p. 2). In other words, it is not as though all linguistics can do is observe people use and

change language in unrestricted ways—there is a constraining topography underlying language use, and studying it can result in linguistic theories that make scientific predictions about possible language use and change.

Turning from the role of grammaticality to the question whether this aspect has become too dominant in generative syntactic research, let us make another comparison: Language use is constrained by grammaticality in the same way that architecture is subject to the laws of statics—a wide range of houses with many different functions can be built, but certain constructions will not hold up even though they could have been useful. While this comparison maintains that grammaticality structures the space of possible language use, it also makes clear that it is only a small portion of what linguistics can and should analyze—just like the laws of statics are an essential part of building a house even though most of the thinking of an architect concerns other aspects of the house and its use. The at times exclusive focus in generative linguistics on an autonomous syntax and grammaticality, emphasizing core language phenomena for a homogeneous community of idealized speakers, arguably has fostered theory-internal debates, abstracting away from the data until the core supports some postulated (but beautiful) principles and assumptions. At the same time, generative linguistics clearly has started moving beyond the glass bead games in the halls of autonomous syntax. While this article is not the forum to give a detailed analysis of these developments, let me just cite two encouraging examples: Firstly, aspects once stipulated in syntax are now recognized to be explainable through the interaction of different modules of linguistic analysis and an explicit analysis of the integration of a sentence into the discourse. The predictions of such theories arise through an interaction of constraints, e.g., connecting syntax, semantics, and information structure (cf., e.g., De Kuthy 2002). Secondly, restrictions once included in the model of syntactic competence have convincingly been argued to result from human sentence processing effects (cf., e.g., Kluender 1998). While both of these developments cut syntax down to size by relocating some of the constraints to other modules of linguistic competence and performance, they keep the central assumption of grammaticality, supporting linguistic theories that make scientific predictions about unseen data.

In conclusion, instead of following Sampson's proposal to throw out grammaticality and the predictive power of linguistic theory, it seems more pro-

ductive to support the reorientation of linguistic analysis towards interacting modules of linguistic competence and performance, and to ground linguistic theorizing empirically by highlighting that a theory is only as good as the empirical insights it provides.

This naturally leads over to the second aspect of Sampson's article to be discussed here, the use and the problems of introspection for empirically grounding linguistic theories. Sampson argues that introspection has no role to play in a scientific enterprise. Instead of native speaker intuitions, he proposes to rely exclusively on "interpersonally-observable evidence"(p. 15). Sampson unfortunately does not list what he considers to fall under interpersonally-observable evidence, but given that the empirical discussion in the paper is based entirely on corpus data (and anecdotes), we can safely assume that Sampson considers corpus data to be a key component in this regard. Using corpus data as evidence for linguistic analysis combines two aspects of linguistic methodology which are useful to make explicit here: On the one hand there is the question how example sentences that are relevant to a particular research question can be obtained. On the other hand one needs to decide whether an example sentence one has obtained is actually part of the language one is studying. How to obtain and how to evaluate example data are two separate questions, which generally require separate answers (cf. Meurers 2005).

The typical method for obtaining example sentences in generative linguistics first considers which properties a relevant example sentence should have and then hand-constructs an example realizing those properties. Such examples usually are not instances of actual language use and typically no context is provided for them. The hand-constructed examples are then evaluated by a native speaker (often the author of the paper) judging whether the sentence is part of their language or not.

In contrast, Sampson in this paper answers the question of how to obtain relevant example sentences by querying an annotated corpus for a set of examples fitting a particular pattern. He then evaluates the example data in two ways: First, he analyzes the obtained examples based on how frequent they are in the corpus. Second, he classifies individual examples by distinguishing "Dunster constructions" from "clearly normal constructions" (p. 8). Sampson characterizes Dunster constructions as "a construction which be-

fore I encountered it I would not have thought of as available in my language, but which after confronting a real-life example I come to see as a valid possibility which had been available to me all along." Nothing is said about how Sampson identifies "clearly normal constructions", but it seems most likely that this judgment is equally based on his intuition as a native speaker of English. Overriding his characterization of intuition as incompatible with a scientific process, Sampson thus relies on speaker intuitions for the evaluation of example data.

An arguably relevant difference between the use of intuition by the generative linguists Sampson discusses and his own use of intuition is the fact that his intuitions are about real-life examples from corpora, whereas the intuitions he discusses as problematic are based on hand-created example sentences. Much can be said in support of using corpus data for linguistic research, and annotated corpora can support searches for sets of examples that are relevant for the construction or validation of linguistic theories (cf., e.g., Meurers 2005). Making explicit why corpora deserve to play a major role as source of examples for theoretical linguistics, in Meurers and Müller (2007) we argue that for studying a linguistic phenomenon "one needs to reduce examples to whatever properties are relevant for the linguistic issues being researched and to vary selected properties in order to explore the grammatical correlations. This is a complex undertaking that assumes an understanding of what properties can play a role for a given linguistic issue—which often is far from clear [...] Corpus data obtained by searching for a linguistically relevant pattern exhibits a wide variation of known and unknown parameters and can include information on the context, as needed for exploring the interaction of constraints from syntax and formal pragmatics. When searching for a particular pattern in a corpus, it thus is possible to observe the theoretically interesting pattern within sentences that exhibit a wide variation of lexical, syntactic, semantic, and contextual properties; this makes it possible to obtain a better picture of which of these properties are relevant for a given phenomenon. The fact that corpus examples generally are natural and contextualized can also be helpful whenever examples are to be evaluated through introspection."

The limits of obtaining data for theoretical linguistic research from corpora should not be overlooked though. Corpora are finite representations of language use and as such can never contain evidence for or against all linguistic generalizations one might want to test. Given Zipf's law that the frequency of

use of the nth most frequently used word or other phenomenon in a corpus is inversely proportional to n, even the largest corpus will appear small for linguistic research looking for specific patterns and theoretical predictions. We here should also remember the point argued above that corpora are merely a source of examples—they do not eliminate the need to evaluate whether a specific example obtained is representative of the language to be analyzed.

Turning to other methods for evaluating example data, the quantitative analysis of corpus data for theoretical linguistics is another rarely used option. Stefanowitsch (2005) illustrates that the common rejection of quantitative analysis in generative circles is based on an impoverished notion of quantitative analysis. An analysis comparing the relative frequency of constructions, taking into account the frequency of its parts and the mode of construction, can support linguistic judgments which are much more sophisticated and linguistically insightful than those built on analyzing bare frequency alone.

Turning to other methods for evaluating linguistic judgments about example data, it is somewhat surprising that Sampson does not mention the well established and validated experimental methods which have been developed over the past decade. As already highlighted by Vasishth (2003), Cowart (1997) and others have established standard procedures for designing syntactic experiments which address Sampson's concerns about unreliable intuitions. Such controlled experimental setups have been shown to provide reliable grammaticality judgments from random samples of native speaker subjects. Arguably this is just the tip of the iceberg of methods which in recent years have become available for empirically grounding theoretical linguistic research. For example, psycholinguistic production and perception experiments, neurolinguistic imaging techniques, or the analysis of transfer in second language acquisition all provide empirical evidence characterizing the set of linguistics categories and forms to be modeled by linguistic theories. They constitute examples of "interpersonally-observable evidence" and some directly address Sampson's concerns about introspection. But none of these methods are mentioned in Sampson's article, presumably because they are not as such dependent on corpus data.

Arguably the biggest mistake of generative linguistics was to throw out most sources of empirical evidence, from rejecting the use of corpora to downplaying the relevance of historical linguistic evidence or observations about

variation in language use. Replacing the narrow generative perspective with a view that is narrowly corpus based, in which the only evidence considered is the language which happens to be used in a given corpus, would amount to replacing one half-blind extreme with another. Any single methodology can mislead, overemphasize some aspects and hide other aspects from view. So while the use of corpus data has much to contribute towards empirically grounding theoretical linguistic research, the key to a successful linguistic science lies in a plurality of methods. Mountain regions are best explored with all our senses.

References

Cowart, Wayne (1997). *Experimental Syntax: Applying Objective Methods to Sentence Judgments*. Thousand Oaks, CA: Sage Publications.

De Kuthy, Kordula (2002). *Discontinuous NPs in German — A Case Study of the Interaction of Syntax, Semantics and Pragmatics*. Stanford, CA: CSLI Publications.

Hock, Hans Henrich (1991). *Principles of historical linguistics*. Berlin, New York, Amsterdam: Mouton de Gruyter, second edn.

Kluender, Robert (1998). On the Distinction between Strong and Weak Islands: A Processing Perspective. In Peter Culicover and Louie McNally (eds.), *Syntax and Semantic, Volume 29: The Limits of Syntax*, San Diego, London: Academic Press, pp. 241–279.

Meurers, Detmar and Stefan Müller (2007). Corpora and Syntax (Article 44). In Anke Lüdeling and Merja Kytö (eds.), *Corpus Linguistics. An International Handbook*, Berlin: Mouton de Gruyter, Handbooks of Linguistics and Communication Science. `http://ling.osu.edu/~dm/papers/meurers-mueller-07.html`.

Meurers, Walt Detmar (2005). On the use of electronic corpora for theoretical linguistics. Case studies from the syntax of German. *Lingua* 115(11), 1619–1639. `http://ling.osu.edu/~dm/papers/meurers-03.html`.

Sampson, Geoffrey (2000). The SUSANNE Corpus: Documentation. `http://www.grsampson.et/SueDoc.html`. Release 5, 2000.08.11.

Stefanowitsch, Anatol (2005). New York, Dayton (Ohio), and the raw frequency fallacy. *Corpus Linguistics and Linguistic Theory* 1(2), 295–301.

Vasishth, Shravan (2003). Review of "Geoffrey Sampson: Empirical Linguistics". *Journal of Linguistics* 39, 688–694.