D.T. Langendoen and P.M. Postal: The Vastness of Natural Languages /Basil Blackwell, London, 1984/

The book to be reviewed — as its title suggests — argues that the collection of sentences comprising each individual NL /natural language/ is vaster than is currently universally assumed. The word 'vastness' is used because, according to the authors, the magnitude of sentences contituting a particular NL cannot be given by any number, finite or transfinite. It is well-known that Turing machines, algorithms, procedures or grammars can construct recursively enumerable, hence countable, collections only. Langendoen and Postal say that the magnitude of sentences of any NL is greater than enumerable infinite, moreover, greater than anybody might have ever thought. Therefore the inadequacy of formal tools listed above can be easily seen.

The authors rightly suppose the reader to be unfamiliar with set-theoretical theorems and the world of paradoxes. To help to understand the book Chapter 1 provides a brief outline of basic set theory. The heart of this chapter is a more or less informal introduction to Cantor's Theorem giving the basis of the later analogy between sets of sets and sets of sentences.

Identification of sentences with sets is known from Johnson and Postal /1980/. We meet the continuation of this idea when Langendoen and Postal apply one of the most important theorems of set theory. In order to use this theorem we have to allow for the existence of sentences with $\aleph_0$ elements /that is, sets with cardinality $\aleph_0$/. Therefore the first part of the book /to support the applicability of Cantor's results/ makes attempts to justify the statement that no arguments can be given against the position that NLs have sentences with transfinite length.

As Matthews /1979/ points out, "a grammar is a partial description of a language; tell us, in certain terms and within certain limits, what the language is like. One way of telling us is to characterize a set of sentences." Chomsky /1959/ chose this way when he outlined the nature of grammars. He gave a definition, divided in Chapter 2 by Langendoen and Postal into the following statements:

"/a/ An NL, e.g. English, can be regarded as a collection of sentences, taking these as formal objects of some type.

/b/ The collections referenced in /a/ are sets.

/c/ These sets are countably infinite.

/d/ Each sentence is a finite object.

/e/ Each NL involves a finite vocabulary.

/f/ For every X, if X is an infinite set, then X can be characterized /'studied'/ only via finite devices that enumerate X. In the case of an NL, the finite device in question is the grammar of that NL." /p. 15/

All of the counter-arguments in this book arise against this definition and no passages speak against the 'ill-defined'-ness of basic notions of this formal-like system.

We cite Matthews /1979/ again: "In formal linguistics a system is well defined. /.../ The set of real sentences is not well defined. /.../ The formal language is a model [of an NL] — a well defined model, albeit of an ill-defined object. We say that such and such grammar generates the sentences of English. But this is not strictly so. In fact, it generates a set of formal sentences then interpreted as a model of what we called the real sentences."

It is obvious that Chomsky has formulated his opinion about NLs with the help of formal languages. But as Moore and Carling /1982/ remark: "It is important to keep in mind that at the outset at least Chomsky was not, in constructing a synthesizing device, attempting to model language in its entirety but rather the syntactic structures underlying the sentences of NL." /p. 55-56/ And these syntactic structures have always been finite, as /d/ shows.

Now, let us see what Langendoen and Postal say about Chomsky's definition. They are in absolute agreement with /a/[1] which serves as a precondition of /b/ and /c/. The crucial point, of course, is the rejection of /d/. Hence, the authors come to the conclusion that the cardinality of any NL is greater than any cardinal number. That is: /c/ is untenable, and what is more, /b/ is either. Furthermore, a description can be given for X without any device enumerating X, hence the assertion that /f/ is false is justified by the existence of non-constructive grammars /see below/.

/e/ is the only statement not directly referred to: if no size laws are valid for any entities of NLs, the vocabulary of morphemes must be infinite. Moreover, there is "no finite bound on the number of phonemes in a morphophonemic representation of a single morpheme". /p. 107/ Independently of the size of the vo-

cabulary, morphemes not occurring in it should be handled in some way. We think that the size of vocabularies cannot be important, if they are open and no syntax is based on them. It is possible to construct grammars having such properties /see Kálmán and Prószéky, 1984a, 1984b/.

Chapter 4, demonstrating the analogy with Cantor's results, is the core of the book. In order to prove that NLs are not sets, the analogue of the notion of power sets is required. Since sentences have been introduced as sets, it is only the power set that has to be given linguistic relevance. Langendoen and Postal define formally the Coordination Projection of a set: every sentence-set univocally defines a compound sentence consisting of clauses contained in the original sentence-set as elements. Thus, the NL Vastness Theorem can now be constructed: NLs are not sets /are megacollections/. The proof is indirect: assume that the language is a set. Then — since NLs are closed under coordinate compounding —, on the one hand, the sets of compound sentences form a subset of the whole language /hence cardinality of the language is greater than their cardinality/ and, on the other hand, the cardinality of the set of all Coordinate Projections of a given language /the power set/ is greater than the cardinality of the language. Since the conclusions contradict each other, NLs cannot be sets. Here we can see the usefulness of Chapter 3: in order to make use of Cantor's Paradix, at least $\aleph_0$ long conjunctions of finite sentences are needed.

The first corollary of the NL Vastness Theorem is the No Upper Bound Theorem: $/\forall k/\ /\text{Cardinal}/k/ \rightarrow /\exists X/\ /X \in \mathcal{L} \wedge \text{Length}/X/ > k//$. Nobody would know what to do with such a theorem in practice, therefore the authors have developed their criticism about the existence of any size laws in the opening sections of the book. This is not without results, because the No Upper Bound Theorem seems to be familiar to the reader, although it is not easy to imagine sentences with a length, for instance, of $\aleph_7$. And the only thing they allowed is the existence of sentences with length $\aleph_0$!

At the end of the chapter a misleading argument is forwarded, which pertains to the assumed absence of any size laws. According to the authors, the proof of the NL Vastness Theorem "does not assume there are transfinite sentences but displays a principle, the closure principle, based on finite sentences, which entails there are transfinite sentences, in fact, sentences of every car-

dinality." /p. 70/ This last statement — in this form — does not seem to be correct. Sentences of the 'original' sets are finite; the power set of a set with cardinality $\aleph_0$ is, however, only greater than the 'original' set if conjunctions with length $\aleph_0$ are allowed as elements. That is, transfinite sentences are introduced by the authors /who consider the Closure Principle for Coordinate Compounding to be a universal linguistic law/, so they are unexcludable elements of a model but not entailed as proper grammatical corollaries.

Chapter 5 deals with the most important implications of the NL Vastness Theorem. A consequence is the NL Non-Constructivity Theorem: no NL has any constructive grammar. 'Constructive' means 'generative' or 'proof-theoretic' here. There follows a list of frameworks falsified as theories of NLs by the NL Non-Constructivity Theorem, which might have been omitted just as well. 27 famous frameworks are mentioned here, and finally the reader is confronted with the following statement: "Any constructive system distinct from all of the preceding frameworks" must be false. /p. 73-74/ All the frameworks listed seem to be treated in an off-hand manner, although the authors mention that the NL Non-Constructivity Theorem does not show that everything about them is incorrect.

Later in this chapter Langendoen and Postal bring a historical parallel: they compare the generative philosophy of Chomsky with their own results. Their claim is that they discovered the same relation between Hilbert's formalist philosophy and Gödel's incompleteness theorems. This parallel is, of course, "based not on the formal character of the proofs but on the way they show the unrelizability of certain intellectual programs" /p. 79/. Comparing mathematics and linguistical modelling is dangerous. Mathematics is a pure, independent science, and linguists when modelling can at best apply /for their ill-defined objects/ the results of mathematics. It takes a lot of courage to call applications independent results.

The formal definition of NCGs /Non-Constructive Grammars/ first appears in Chapter 5.2 of this book:[2]

$$/\forall X/\ /NL/X/ \leftrightarrow X \subseteq US \wedge COLA/X/ \wedge /\exists Y/\ /Grammar/Y/ \wedge Y =$$
$$= \{ SR_i, CR_j \} \wedge CR_j/X/ \wedge //\forall z/\ /z \in X \rightarrow SR_i/z/////\ ,$$

where US is a subset of all pair networks[3] for which all of the sentence laws are true; SR_i and CR_j are some proper logical conjunctions of sentence rules and corpus rules, respectively. Corpus

rules are NL-specific, corpus laws /COLA/ are universal. In particular, being closed under coordinate compounding is a corpus law.

An NCG consists of a finite set of statements to which truth values can be assigned. The set of these statements picks out a subset of US. Readers unfamiliar with Johnson and Postal /1980/ feel the need for an example to make the NCGs easier to understand, but they cannot find any examples at all in this book.

In Chapter 5.3 the question of effability still comes up in connection with Katz, who has argued that "each proposition /thought/ is expressible by some sentence in every NL." /Katz, 1981, p. 226/ According to Russell /1903/ there is a megacollection of propositions. Langendoen and Postal consider as a positive consequence of the NL Vastness Theorem that it keeps the effability principle consistent with Russell's claim.

The core of Chapter 5.4, about language learning, is a theorem--formulated carelessly, the Psychogrammar Theorem: "Let P be a generative psychogrammar of L and let J be an NL. Then: $J \neq L$." /p. 98/ One could ask why J and L should be equal in the first place. But let us ask the less demanding question why one should deal, in a book about mathematical linguistics, with a notion even more ill--formed than the other notions occurring in the book? Following Bever, Langendoen and Postal say that P is a psychogrammar of an NL L iff the range of P includes all the sentences of L known by a speaker of L. The conclusion of this chapter is that not every NL is learnable /p. 102/, because of limitation on time, computing space and storage place. In spite of this, according to the authors, inquiry into the properties of finite sentences is surely of psychological relevance, but it "should not be confused with grammatical study proper, any more than size-bounded computer arithmetic should be confused with the principles of number theory." /p. 103/ This analogy is, however, drawn incorrectly. The way how transfinite sentences have become important in this book should rather be compared with constructing cardinal numbers patterned after natural numbers. People use natural numbers in everyday life, and mathematicians have extended the notion of natural numbers in order to characterize infinities. Nobody says that people count with cardinal numbers when they want to know the quantity of something, but they use a subset of the cardinal numbers only, because of the limitations on time, computing and storage space. This subset is the set of finite cardinal numbers, that is, the

natural numbers. 'Study of counting' must not deal with transfinite numbers, although there can exist a theory these numbers play an important role in. Similarly, we can say that the study of grammar must not deal with transfinite sentences /because of its definition, e.g. in Chomsky, 1959, not, however, because of the existence of any size law/, although there is a theory transfinite sentences play an important role in. This theory is based upon a set-theoretical model of language. It would not seem surprising that Chapter 7 shows us that "the only properties systematically distinguishing transfinite from finite sentences of a fixed NL are those which in general distinguish transfinite from finite sets." /p. 166/

We do not intend to enter into details in connection with the philosophical discussions of Chapter 6, although this is the largest chapter of the book. One can easily accomodate different answers to the question about the ontological status of universals. The main traditional answers are known as realism, nominalism, and conceptualism. Platonic realism holds that universals ar real, but distinct from mental or physical objects. Nominalism holds that only the sensible signs of language are real. Conceptualism holds that universals are mental. This chapter deals first of all with three particular versions of the conceptualist position /standard, performance and radical conceptualism/ and their relation to transfinite sentences. The conclusion of this chapter is that "there are no grounds in ontological commitments which would justify the claim that coordinate compounds with infinitely many conjuncts lie beyond the bounds of proper linguistic characterization." /p. 156/

The notion of NCGs — as we noted earlier — arose for the first time in Johnson and Postal /1980/, although other NCGs have been developed since then /see Kálmán and Prószéky, 1984a, 1984b/. At that time the characterizability of transfinite sentences was not mentioned among their advantages, probably because it had not been required by any phenomena discussed.

The real benefit of Langendoen and Postal's book is raising the possibility of NCGs and verifying their competitiveness with the well-known contemporary models, first and above all the generative one. According to this, an NCG is a formal device not synthesizing and analyzing anything. It is a descriptive model only, it makes statements about sentence structure, but not generating or parsing rules. Such kinds of assertions can characterize even elements without any size limitations. But this feature

of NCG must not be considered fundamental, because NCGs seem to be the only possible way to replace some questionable viewpoints established by the generative grammars. People reading the book by Langendoen and Postal will, however, consider not the real possibilities, but the describability of transfinite sentences as the most essential /if not the only/ advantage of NCGs.

It is important to specify that transfinite sentences are contained not by NLs, but by a model of NLs allowing infinite conjunctions. To allow things like this has never come to Chomsky's mind: his language model relying on finite sentences has remained unquestioned up to now from a mathematical point of view, at least for the finiteness of sentences. Hence no mathematician would require transfinite sentences to occur in the language model. On the other hand, the job of linguists has not been hindered by the least upper size bound occurring /implicitly/ in the usual definitions. Thus, featuring of transfinite sentences is solely a consequence of a particular model, but not of NCGs generally.

The source of the problems discussed in this book is the well-known view that an NL must be considered as a collection /of strings, sets or whatever/. However, the real advantage of the NCGs lies in the fact that there can exist grammars not referring to any collection and still being able to characterize languages by virtue of assertions about the form of utterances. Grammars of this sort describe, in fact, how the set of elements becomes the structure of a sentence. In order to construct or identify the syntactic role of an element we do not need any vocabularies, finite or infinite. This can easily be seen in the case when unknown elements occur in an utterance or known elements are in 'strange' roles /see Kálmán and Prószéky, 1984a, 1984b/. The possibility to handle these phenomena has its roots in referring to no collections at all. In other words, the collection of utterances is open. Consequently, the use of NCGs might, and does, allow the distinction between immobile linguistic description and the strategy to make them function.

These possibilities can be realized only if, much in contrast to Langendoen and Postal, we take the most important properties of NCGs to reside not in the possibility to characterize transfinite sentences. This could only be expected of a book dealing first of all with language and not the outcomes of modelling with the help of set theory.

Notes

/1/ We are not. See Kálmán and Prószéky /1984a, 1984b/ and this review further below.

/2/ See, for its definition, Johnson and Postal /1980/.

/3/ Their definition is given in Johnson and Postal /1980/.

References

Chomsky, N. /1959/ 'On Certain Formal Properties of Grammars'. Information and Control, 2:137-167.

Johnson, D. and P. Postal /1980/ Arc Pair Grammar. Princeton University Press, Princeton, New Jersey.

Kálmán, L. and G. Prószéky /1984a/ 'FMR Grammar'. /In Hungarian/ Magyar Nyelvőr, forthcoming.

Kálmán, L. and G. Prószéky /1984b/ 'Against Constructive Grammars: Analogy and Productivity in the FMR Grammar'. Computational Linguistics, Logic, Languages, in preparation.

Katz, J. /1981/ Language and Other Abstract Objects. Rowman and Littlefield, Totowa, New Jersey/Basil Blackwell, Oxford.

Matthews, P. /1979/ Generative Grammar and Linguistic Competence. George Allen and Unwin, London.

Moore, T. and C. Carling /1982/ Understanding Language: Towards A Post-Chomskyan Linguistics, Macmillan Press, London.

Postal, P. /1982/ 'Some Arc Pair Grammar Descriptions'. in P. Jacobson and G. Pullum /Eds./ The Nature of Syntactic Representation. Reidel and Company, Dordrecht.

Russell, B. /1903/ The Principles of Mathematics. Norton and Company, New York.

Gábor Prószéky
Institute for Musicology
Hungarian Academy of Sciences
Budapest
Táncsics M. u. 7.
Hungary
H-1014