Esa Itkonen, University of Turku, Finland

Comment on Klaas Willem's contribution

Willems approaches one of the fundamental questions of linguistics: Is linguistic data ultimately discrete ('all-or-none') or continuous ('more-or-less')? The rise of phonology originally convinced many people of the correctness of the former alternative:

> […] it was considered a decisive accomplishment to show the existence of qualitative structure in spheres of human life (cf. Whorf 1941). The basic units of phonology and morphology were salient exhibits of this cause. Linguistics was a demonstration of the possibility of rigorous formal analysis of a sort not requiring sampling, statistics, or other techniques derivative of a natural science orientation […] (Hymes & Fought 1981: 175; emphasis added).

On the other hand, Bolinger (1961), for instance, pointed out that there are many areas of linguistic research where the discreteness thesis has to be, and is, abandoned. Systematic variation is a prime example, of course. But gradients or continua have to be established already at the conceptual level. For instance, the noun vs. verb distinction may seem discrete, but there are (at least in the IE languages) also a few intermediate members, which means that we have to do with a continuum: noun < nominalization < participle < infinitive < finite verb.

The example that Willems adduces, i.e. the German verb aufsetzen + auf + ACC/DAT, seems clear enough. In today's language the auf + ACC construction is no longer a genuine option. It is only the conservative (= contrary-to-fact) intuition of the grammarian which still maintains it. For my part, I would like to add the following example.

In English, for instance, complementation may be expressed either by that-clauses or by non-finite constructions:

(1)   John knew that the children were asleep
(2)   John knew the children to be asleep
(3)   John saw that the children were coming home
(4)   John saw the children coming home

In Finnish, there is no direct counterpart to (2). Rather, non-finite complementation is uniformly expressed by a participial construction similar to (4). In order to follow my argument, it is enough to know that (1) and (2) are translated into Finnish by (5) and (6), where että = 'that', and nukkuivat and nukkuvan are finite and non-finite, respectively:

(5)   John tiesi, että lapset nukkuivat
(6)   John tiesi lasten nukkuvan

In what follows, (5) and (6) will simply stand for the two options. Now, Hakulinen and Karlsson (1979: 356) make the following claim, adapted to our examples (5)–(6): "In

complementation, (5) represents the norm. Some verbs allow (5) and exclude (6), while other verbs allow both (5) and (6). When this is the case, (5) and (6) are in free variation."

This sweeping claim is simply based on the linguistic intuition of the two authors, and it is rather plausible as long as single example sentences analogous to (5) and (6) are considered in isolation. If, however, corpus data are taken into consideration, then – as has been shown by Pajunen (2001: 378–412), based on a corpus of 24 million words – every component of this claim turns out to be false:

(i)   Each of the verbs which are claimed to exclude (6) in fact allows (6).
(ii)   With some of the verbs claimed to exclude (6) in favor of (5), (6) is actually more frequent than (5).
(iii) There are some verbs which, contrary to the claim, exclude (5) and allow (6).
(iv)   When both (5) and (6) are allowed, they are never if free variation.
(v)   The difference between (5) and (6) is conditioned by the following factors: a) the meaning of the governing verb: speech act vs. cognition vs. perception verb, b) ACT vs. PASS verb forms of the governing verb, c) same-subject vs. different-subject constructions (in ACT verb forms), d) the simultaneity vs. non-simultaneity of the events spoken about by the two verbs.

This argument is reproduced in Itkonen & Pajunen (2010: 61–70) as part of Chapter 2, which deals with the methodology of corpus analysis.

In brief, we have encountered here the limits of intuition: in those ('less-than-clear') cases where intuition is not enough, one has to resort to the use of observational or experimental evidence.

It is crucially important to understand that the existence of less-than-clear cases by no means invalidates the existence of clear cases. These are two different things, even if the actual dividing line necessarily remains vague. My own favorite example of clear cases is provided by those 39 either grammatical or ungrammatical sentences, given in Chomsky (1957), which constitute the data basis of the 'generative revolution'. Of course, this data basis is surprisingly scarce. But at least it is secure: no one has ever contested or will ever contest the grammaticalness of John has read the book; nor has anyone contested or will ever contest the ungrammaticalness of *Of admires John. In other words, these are clear cases, where intuition is enough, i.e. where "no one has yet found any disagreements that would move us to begin a program of observation and experiment" (Labov 1975: 8).

In what precedes, I have outlined my solution to the fundamental problem of "how observation and experiment are to relate to the intuitive data in the actual work of linguistic analysis" (op. cit., p.54). It follows inter alia that it makes no sense at all to claim, as Newmeyer does, that "grammar is grammar and usage is usage".

In my own contribution I illustrate the notion of iconicity with the SG vs. PL distinction: what is linguistically less vs. more corresponds to, and is explained by, what is ontologically less vs. more. Therefore it seems appropriate for me to conclude with a comment on Haspelmath's claim (discussed also by Willems) that frequency is enough to explain what others (including myself) claim to be explained by iconicity.

The counter-argument against Haspelmath refers to reduplication as the (iconic) expression of semantic/ontological plurality. Here are a few examples taken from Itkonen (2005).

Diyari: In addition to dedicated markers of dual (-wulu) and plural (-wara), there is also a more or less nonsystematic use of reduplication, e.g. kupa ('child') vs. kupa-kupa ('children') (pp. 26–27).

Hua: In addition to other PL markers, there is a systematic use of reduplication in nouns ending in -V': e.g. eva' ('rock') vs. eva'-ava' ('rocks'), huga' ('frog') vs. huga'-aga' ('frogs') (p. 68).

Rapanui: In intransitive verbs, reduplication of the first syllable expresses plurality of subjects whereas in transitive verbs it expresses plurality of objects (p. 121).

Wari: Throughout the grammar, plurality is expressed by reduplication, starting with mao na ('s/he went') vs. mama nana ('they went') (p.168, and passim).

It is remotely possible to assume, with Haspelmath, that developments like Latin amicus (NOM.SG) vs. amici (NOM.PL)/amicos (ACC.PL) > Spanish amigo (SG) vs. amigos (PL) supports the view that the shorter SG form is due to its greater frequency vis-à-vis the corresponding PL form. Hence, the SG form would result from some kind of erosion. But it is not even possible to form a coherent idea of what kind of 'erosion' would produce the Hua SG form eva' as opposed to the PL form eva'-ava'. Therefore the only coherent alternative is to accept explanation-by-iconicity, or – more generally – the existence of a "pervasive intention to produce shorter forms for particular functions", as Willems puts it.

References

Bolinger, Dwight (1961): *Discreteness, Gradance, and All-or-None*. The Hague: Mouton.

Chomsky, Noam (1957): *Syntactic Structures*. The Hague: Mouton.

Hakulinen, Auli & Fred Karlsson (1979): *Suomen kielen lauseoppia* ('On Finnish Syntax', in Finnish). Helsinki: Suomalaisen Kirjallisuuden Seura.

Hymes, Dell & John Fought (1981): *American Structuralism*. The Hague: Mouton.

Itkonen, Esa (2005): *Ten non-European Languages: An Aid to the Typologist*. University of Turku: Publications in General Linguistics 9.

Itkonen, Esa & Anneli Pajunen (2010): *Empiirisen kielitieteen metodologia* ('The Methodology of Empirical linguistics'). Helsinki: Suomalaisen Kirjallisuuden Seura.

Labov, William: What is a linguistic fact? Lisse: Peter de Ridder Press.

Pajunen, Anneli (2002): *Argumenttirakenne* ('Argument Structure', in Finnish). Helsinki: Suomalaisen Kirjallisuuden Seura.

Willems, Klaas (2013): Is Frequency an Explanatory Causal Concept in Linguistics? Energeia 5.