

Excerpts from 'Comments on *World Atlas of Language Structures*' by Esa Itkonen & Anneli Pajunen

(1) The two categories expressed by the most synthetic verb form of English are said to be "agreement (in the present: -s) and tense (past: -ed)" From this it is concluded that "[t]he English verb therefore has a synthesis degree of 2 cpw [= category-per-word]".

This is very odd. The affixes -s and -ed never occur in one and the same (syntagmatic) wrd. If **the** most synthetic **verb form** of English has the value of 2 cpw, we have the right to know what it is. But there is no such thing. Instead, there are **two** verb forms, e.g. *smoke-s* and *smoke-ed*. These are (paradigmatic) **alternatives**. (There is no such verb form as **smok-ed-s*.) This means that synthesis, as measured by the cpw value, is now defined as the number of all grammatical meanings that can be expressed by affixes/clitics attached to different verb forms in a given language. (Thus, to repeat, the syntagmatic and the paradigmatic dimensions are being confused here.) This is the **third** interpretation of 'synthesis'. It is clear, however, that if terms are used in their proper meanings, then English has the value of exactly 1 cpw: either *smoke-s* or *smoke-d*.

To fix our ideas, let us compare English with Finnish. If measured by the maximal number of affixes/clitics in one and the same (syntagmatic) verb form, Finnish seems to have the value of 4 cpw, as in *ost-i-mme-ko-han* (= buy-PAST-1PL-Q-CL = 'did we buy, I wonder?'). If we adopt the (quixotic) third interpretation, then the Finnish cpw value rises to 6: (subject) agreement, (past) tense, (irrealis) mood, (passive) voice, (interrogative) illocution, discourse 'coloration'. (And the number rises even further if we take into account such categories as causativity and affirmative polarity.) Where Finnish has the values 4 and 6, English (as shown above) has the values 1 and 2.

The scale of synthesis can be summarized as follows:

analytic < synthetic
synthesis: low (= isolating) < medium < high (= polysynthesis)

On the map (pp. 97-98) the two extremes of this scale are exemplified, as expected, by isolating languages (like Mandarin Chinese), on the one hand, and polysynthetic languages (like Wichita), on the other. But otherwise the resulting cpw values are often incomprehensible. Here are a few selected oddities:

... < 2-3: English, Finnish < 4-5: West Greenlandic < 6-7: Yoruba < 8-9: Rapanui < 10-11: Yagua

It was seen above that it is just wrong to consider English and Finnish as equally synthetic. West Greenlandic is a language of "high polysynthesis" (Mithun 1999: 407): For instance, Fortescue (1984) gives several examples of verbs containing 8-9 affixes expressing subject and object agreement, tense, aspect, voice, causativity, polarity, etc. Because, in connection with fusion, Yoruba was claimed to be the only (!) 'tonal/isolating' language, it makes no sense to claim that it is strongly synthetic, and even more so than (polysynthetic) West Greenlandic. The Rapanui verb has no agreement, and its TAM system is expressed by clitic **frames**, which means that they exemplify the 'extended' or 1M2F type. (This is how most of the 'positions' in a verbal frame have to be interpreted; cf. Du Feu 1996:12.) The cpw value of 8-9 is simply impossible. The Yagua verb is able to contain clitics for agent, patient, and recipient as well as affixes for tense, aspect, causativity, and motion. Still, Payne & Payne (1990) give no examples of verbs containing more than 6 affixes.

(2) Finnish and Nenets are wrongly classified as having (cumulative) case & number, and on this very slender (and mistaken) basis IE languages and Uralic languages are assumed to have descended from a common protolanguage. In fact, the Uralic protolanguage already had the “non-singular” marking *-j reflected e.g. in the Finnish *talo-issa* (‘house-PL-INESSIVE’ = ‘in [the] houses’).

(3) If A and P are treated differently in a language, it is the treatment of P which is chosen as representative. This decision has the unwelcome consequence that a language with dependent marking plus subject agreement but without object agreement is assigned the same ‘locus type’ as a language with dependent marking but without subject and object agreement. Hence Finnish, for instance, ends up having the same type as Japanese and Diyari (i.e. ‘dependent-marking’). Here vital information remains hidden (if not deliberately concealed).

(4) Chapter 24 *Locus of Marking in Possessive Noun Phrases*: Finnish and Hua are mistakenly labelled ‘double marking’. In the case of Finnish, there is probably a confusion with the dependent-marking nominal possession (*miehe-n sika* = ‘man-GEN pig’ = ‘the man’s pig’) and the double-marking pronominal possession (*hän-en sika-nsa* = 3SG.HUMAN-GEN pig-3SG/PL.HUMAN.POSS = ‘his/her pig’). In the case of Hua, there is probably a confusion between the dependent-marking nominal possession (*de-ma’ fu* = ‘man-GEN.SG pig’ = ‘the man’s pig’) and the head-marking pronominal possession (*fu-’a* = ‘pig-3SG.POSS = ‘his/her pig’) (cf. Haiman 1980: 238-239, 224-225).

(5) In Chapter 30 *Number of Genders* the term ‘gender’ is taken to be synonymous with ‘noun class’. The most important gender divisions are ‘male vs. female (vs. other)’, ‘human vs. non-human’, and ‘animate vs. inanimate’.

Although English nouns lack gender, English is nevertheless deemed to possess a pronominal gender system, because of the **agreeing** personal pronouns *he, she, it*. Such systems are claimed to be “rare”. This is not quite right, however. Finnish, for instance, has the same type of system, based on the ‘human vs. non-human’ distinction: 3SG = *hän* (‘he/she’) vs. *se* (‘it’), 3PL = *he* (‘they.HUMAN’) vs. *ne* (‘they.NON-HUMAN’). This distinction is accentuated by a further syntactic distinction: *hän/he* have possessive inflection while *se/ne* have not.

Nevertheless, Finnish is classified as having no gender. The reason must be that while *hän/he* is a personal pronoun, *se/ne* is taken to be a demonstrative pronoun, more precisely an “anaphoric-neutral pronoun” (Abondolo 1998: 170). But here the diachronic-structural point of view is given primacy over the synchronic-functional one, which is wrong. Functionally, *se/ne* is the exact non-human counterpart of *hän/he*. Let us assume that assertions with the following meanings have been uttered in Finnish: ‘There was my brother ~ sister vs. my dog ~ book’. Then it would be normal to add the questions: *Näitkö hänet vs. sen?* (= ‘Did you see *hän-ACC* vs. *se-ACC*?). Analogous PL forms of *he* vs. *ne* would be anaphorically used for ‘my brothers ~ sisters vs. dogs ~ books’. Thus, the criterion of agreement is fulfilled. If English has a pronominal gender system, so has Finnish.

A further objection would refer to the fact that *se/ne* are not **dedicated** morphemes, i.e. they have other meanings besides ‘3SG/PL.NON-HUMAN’ (and in the spoken language they are supplanting *hän* vs. *he*). But this just shows the questionable nature of the very notion of ‘dedication’. There will be reason to return to this point.

Let just add one more point. What is so special about (agreeing) **personal** pronouns? If **interrogative** pronouns are accepted as well, we see in Finnish the same ‘HUMAN vs. NON-HUMAN’ distinction, expressed by the formal distinction *ku-ka* vs. *mi-kä* that goes back to the Uralic protolanguage and recurs in many Uralic languages today. Of course, this distinction is nearly universal in the world’s languages, and it certainly exhibits agreement

phenomena of one type or another. So there may be reason to rethink the notion of gender to some extent.

Finnish is classified as having no gender also in Chapter 44 *Gender Distinctions in Independent Personal Pronouns* (p. 184), although the *hän/he* vs. *se/ne* distinction clearly satisfies the value “gender distinctions in third person only, but in both singular and nonsingular”, which is claimed to occur (only) in 11 per cent of the world’s languages. Therefore, in light of what precedes, the following generalization may well be questioned: “the main areas **without** gender are ... Siberia (notably the families grouped under Uralic and Altaic)” (p. 127).

(6) Optionality should have been one of the parameters in Chapters 37 *Definite Articles* and 38 *Indefinite Articles* (as it was in Chapter 34). Now it is just informally noted that one language (e.g. Tauya) has an optional indefinite article but no definite article whereas another (e.g. Kutenai) has an optional definite article but no indefinite article; but these facts do not show up in the statistics. Abandoning the discrete either-or approach would have made it possible to record many facts of the same kind. For instance, prenominal definite and indefinite articles (i.e. *se* and *yksi*) are emerging (or have already emerged) in spoken Finnish; and Yoruba has optional postnominal indefinite and definite articles (i.e. *kan* and *náà*).

Mansi is classified as having no definite (or indefinite) article, but in fact, 3SG forms of possessive declension have come to signify definiteness (cf. Keresztes 1998: 417). Every now and then we all need to be reminded of the basic fact that languages change: a form or construction ‘dedicated’ to one function (e.g. possession) may become ‘dedicated’ to another function (e.g. definiteness) or may become non-dedicated, i.e. polysemous, so as to have two functions (e.g. possession and definiteness), or even more.

(7) Chapter 50 *Asymmetrical Case Marking* concentrates on whether or not there are differences between case paradigms of various subclasses of nominals (in particular, substantives and personal pronouns). In some languages (like Turkish) there are no such differences. It is rather customary, however, that – in conformity with the empathy hierarchy – the pronominal (case) inflection is more differentiated than the substantival one. A standard example is the English *I* vs. *me* (etc) distinction which has no counterpart among nouns. It is not hard to find analogous instances, starting with the Yoruba 1SG distinction *mo* (direct) – *mi* (oblique, i.e. PAT/REC and GEN), etc. The situation is, however, complicated by the fact that only independent or “free-form pronouns” are taken here into account. This leads to hair-splitting decisions. For instance, Italian and Spanish are supposed to have “asymmetrical case marking” (i.e. **more** case-inflected free pronouns than nouns) while French is not. More precisely, in Italian and Spanish this is asymmetry of the “additive-quantitative” kind. If the minor subclass under examination happens to exhibit **less** cases than the bulk of nouns, then we have asymmetry of the “subtractive-quantitative” kind. This too can lead to difficult decisions concerning languages where both types of asymmetry exist.

In Finnish, personal pronouns have the following case inflections, illustrated here with 1P forms: NOM.SG = *minä* (‘I’) vs. NOM.PL = *me* (‘we’), ACC.SG = *minut* (‘me’) vs. ACC.PL = *meidät* (‘us’), GEN.SG = *minun* (‘my’) vs. *meidän* (‘our’). In comparison, the substantival inflection is defective or syncretic in the sense that the ACC function is performed by forms identical with either GEN or NOM, illustrated here with the word *talo* (‘house’): NOM.SG = *talo* vs. NOM.PL = *talot*, ACC.SG = *talon* (or, with affirmative imperative *talo*) vs. ACC.PL = *talot*, GEN.SG = *talon* vs. GEN.PL = *talo-j-en* (where *j* [*<i>*] is the plural marker). It is presumably because of these facts that Finnish qualifies (p. 208) as “additive-quantitatively asymmetrical”. But there is also another side of the coin: There are three cases, i.e. abessive, comitative, and instrumental (occasionally called ‘instructive’), in

which personal pronouns – unlike nouns – do **not** inflect. Hence, Finnish is **also** “subtractive-quantitatively asymmetrical”. Now, it is decreed that in a case like this the former asymmetry outranks the latter. But why should this be? Why can the two types of asymmetry not be admitted to coexist? Once again, linguistic data are forced into a discrete straitjacket for no real reason.

(8) Chapter 103 *Third-Person Zero of Verbal Person Marking*: In 36 per cent of the languages there is one or another type of zero realization of 3SG/PL. Finnish is wrongly classified (on p. 420) as having “zero in all 3SG forms”. Contrary to this claim, in the PRES.IND paradigm of the Finnish verb, 3SG is typically expressed by lengthening the final vowel (and, often, by internal consonantal change): *laula-n* (‘sing-1SG’), *laula-t* (‘sing-2SG’), *laula-a* (‘sing-3SG’, as opposed to *laula-I* = ‘sing!’); *anna-n* (‘give-1SG’), *anna-t* (‘give-2SG’), *anta-a* (‘give-3SG’). Some 3SG zero forms do occur in this paradigm, but they are the exception.