

ALT 8, Berkeley, Cal., USA, 23–26 July 2009

Morphological cumulation through phonological fusion?

When PERSON is, and remains, separate from NUMBER

Frans Plank, Thomas Mayer, Tikaram Poudel, Michael Spagnol

Universität Konstanz

1. Of all inflectional categories, PERSON – the grammaticalisation of speech-act roles: speaker, addressee, neither-speaker-nor addressee, to only mention the crosslinguistically most common three-way contrast – is more likely to be expressed cumulatively than any others. On a worldwide scale, for all other inflectional categories, both verbal and nominal, it is far more common to be expressed separatively than cumulatively. (It is in fact one family, Indo-European, which contributes the lion's share of cumulation for all other categories.) The obvious question is: Why is PERSON special?

2. The inflectional category most likely to be cumulated with PERSON is NUMBER. There is in fact a question here whether it is really NUMBER, strictly speaking, which is the category we are dealing with here. Especially with 1st and 2nd PERSON, the quantificational category concerned is more appropriately assumed to be ASSOCIATIVE ('and others in some relevant way associated with X'). Also, depending on the formal structure of the paradigmatic system concerned, the quantificational contrast may more appropriately be conceived of as one of AUGMENTED vs. MINIMUM reference. In all variations on this quantificational theme, PERSON must be seen as conceptually independent of the quantificational category: there are two dimensions for independent variation, and it is only a matter of form whether these two dimensions are expressed separately or jointly. (And there is other morphosyntactic evidence strongly supporting this conclusion.) Concepts or meanings as such, thus, cannot be held responsible for the unusual proneness of PERSON and NUMBER (or ASSOCIATION or AUGMENTATION) to cumulation. It would seem that this question continues to be wide open.

3. It is often assumed that morphological cumulation is brought about diachronically through phonological fusion: sandhi processes of various kinds would obscure and eventually obliterate morpheme boundaries separating two or more syntagmatically co-occurring exponents. With fusion assumed to be virtually the only mechanism transforming separation into cumulation, separate terms have sometimes been found redundant: hence, the popular term "fusional" morphology instead of "cumulative" morphology.

To be a bit more specific about this diachronic scenario, it would need four conditions for such transformations to be successful:

- the frequent, if not regular, syntagmatic co-occurrence of exponents;
- the adjacency of these co-occurring exponents – for you don't get fused at a distance;
- active processes of irreversibly fusional phonology;
- phonological forms of the exponents concerned which render them, and the boundaries between them, vulnerable to active processes of fusional phonology.

Now, perhaps the exponents of PERSON and NUMBER or other quantificational categories are unusual insofar as they are proner than all others to meet these conditions for phonological fusion. This explanation would seem a bit implausible; but can it be tested and falsified? We cannot really claim to have done this; but we think we can demonstrate, if only inferentially, that the phonological fusion scenario is a non-starter, or, not to be too harsh, rather an also-ran. The only variable that we have indeed tested for is adjacency. And the argument is indirect: What we find about the adjacency and non-adjacency of PERSON and NUMBER (or, as the case may be, ASSOCIATION or AUGMENTATION) cannot be plausibly squared with what one would expect on a phonological fusion scenario.

4. What we examined are bound exponents of PERSON, affixed or cliticised to verbs, nouns, or adpositions, which are separate from exponents of NUMBER (ASSOCIATION or AUGMENTATION). For present purposes we only included information on separation and cumulation in the case of independent pronouns as a control. For bound as well as independent pronouns, separation is the clear minority pattern across languages, with cumulation much more common; and our sample

of languages is therefore comparatively small, currently approaching a hundred. It is not an elaborately constructed sample either: we included whichever languages we found where PERSON is separative.

Not all bound PERSONS will perforce show the same separative or cumulative behaviour: we therefore coded 1st (exclusive and inclusive), 2nd, 3rd PERSONS individually, in case they were not all separative. Likewise, NUMBERS have to be kept distinct too, since, for example, DUAL may be separate from PERSONS, while PLURAL is cumulated. Furthermore, not all series of bound PERSON markers need to show the same behaviour: subject PERSON may be different from object PERSON, or from possessor PERSON; PERSON in the indicative mood may be different from PERSON in the imperative; and similar differences may obtain between tenses or moods, with, say, PERSON separative in the imperfective and cumulative in the perfective.

A further analytic decision was: What is a PERSON marker? We took a liberal stance and included combined, morphologically unsegmentable markers of two PERSONS, like '1st PERSON acting on 2nd PERSON'; they would be separative when not including NUMBER marking for either PERSON involved. Perhaps more controversially, DIRECT and INVERSE markers (as in the Algonquian languages) were also included: they are a sort of combined markers for two PERSONS, in connection with a referential hierarchy also providing information about the syntactic-semantic relations of the PERSONS combined. (They are thus wholly different from PASSIVE markers, which have no inherent PERSON dimension.)

The most difficult decision of all was to decide when we are dealing with separation and when with cumulation. Needless to say, especially in cases where fusion is in the middle of its work, clearcut decisions are unviable. Reference grammars often do not provide neat morphological segmentations even when they would seem obvious; but do-it-yourself analyses of languages the typologist does not know well are fraught with risks. Our tendency was to only accept as separative the really clear cases, where grammarian and typologist would find themselves in comfortable agreement.

Also, we were not always as thorough as would perhaps be desirable and selected a clear case as an exemplar and did not worry over all these less clear cases in the same language. The exemplars in our survey often are bound PERSON markers representing (agreeing with, cross-referencing) intransitive subjects.

But we won't bore you any further with our methodological troubles. Here is what we

(think we) were looking for and what we (think we) found.

5. With adjacency as the only variable tested for, what could we have found concerning bound PERSONs separate from NUMBER (ASSOCIATION, AUGMENTATION)? These four patterns:

- (I) a. ... -PRS-NMB- ... b. ... -NMB-PRS- ...
(II) a. ... -PRS- ... -NMB- ... b. ... -NMB- ... -PRS- ...

PERSON and NUMBER adjacent, in whichever position relative to their base or host, with PERSON preceding (Ia) or following (Ib) NUMBER. Or PERSON and NUMBER non-adjacent, in whichever position relative to their base or host, again with PERSON either preceding (IIa) or following (IIb) NUMBER.

Did we find all four possible patterns, and did we find them about equally frequently?
No, we didn't.

6. The (b) patterns are exceedingly rare. With far more than chance frequency, bound PERSON precedes bound NUMBER, whether the two are adjacent or non-adjacent.

🍏🍏🍏 Counterexamples? discuss briefly 🍏🍏🍏

We will not here attempt an explanation. We would only like to point to the exceptional behaviour of bound PERSON also with regard to the suffixing/encliticising preference: differing from most and perhaps all other bound inflectional categories, including NUMBER (ASSOCIATION, AUGMENTATION), bound PERSON is also frequently prefixing/procliticising. Perhaps the typical grammaticalisation history of bound PERSON helps to account for these somewhat unusual positional proclivities: bound PERSON markers typically derive from resumptive pronouns in topic-comment constructions ('The queen, she(-)abdicated'), and such resumptive pronouns will tend to come earlier rather than later in the comment part ('?The queen, abdicated(-)she'). They will perhaps not only come before than the stem of the comment predicate, but also before other material bound to that stem.

7. For PERSON-before-NUMBER we found both the adjacent (Ia) and the non-adjacent (IIa) pattern – but the former much more frequently than the latter: With more than chance frequency, bound PERSON(s), when separate from NUMBER(s), is/are adjacent to NUMBER(s).

🍏🍏🍏 Survey 🍏🍏🍏

🍏🍏🍏

8. Any generalisations about which particular (bound) PERSONS are especially prone to be separate from NUMBER? (and to be adjacent or non-adjacent?) 2nd > 1st/3rd? (because 2PL reanalysed as 2SG.FORMAL, and 2PL is newly created by means of a separate PL marker being added)

Which particular series? imperfective > perfective (why?)

🍏🍏🍏

9. That bound PERSON and NUMBER tend to be adjacent rather than non-adjacent does not only hold for our (no doubt imperfect) sample of individual languages: it also holds at the level of families. That is to say, if one member of a family has PERSON(s) separate from NUMBER(s) and the respective exponents are adjacent, the probability is high that other members of the same family will also have adjacent separative PERSON(s) and NUMBER(s). (Which sort of accounts for our somewhat unbalanced sample.) Although differences, as brought about by changes, are possible, families are remarkably homogeneous in this respect. We cannot be very specific about the time-depth of many of the families in our survey, and we may sometimes be comparing incomparables insofar as time-depths differ considerably for individual families. Nonetheless, we would like to tentatively conclude that in the clear predominance of adjacency over non-adjacency of bound PERSON and NUMBER markers we have found a pattern that is remarkably time-stable. Once a bound separative PERSON marker finds itself adjacent to a NUMBER marker, this state of affairs is likely to continue unaltered.

10. What does this mean? Here our reasoning becomes inferential. If phonological fusion were the main agent in transforming separation into cumulation, we would not expect so much adjacency – which is after all the license to fuse – with separation of PERSON from NUMBER. It

should instead be non-adjacency, where the crucial condition for fusion is not met, which is conducive to separation. Something must be wrong, therefore, with the fusion scenario; it does, at any rate, not account for the cumulation preference of bound PERSON.

11. If we find so many instances of bound PERSON and NUMBER markers which appear to have remained unfused over some possibly considerable time even though they could have fused, co-occurring regularly and being adjacent, the fault might also lie with the other preconditions for fusion: the existence of active processes of irreversibly fusional phonology in the families concerned, and phonological shapes of the markers concerned which would render the boundaries between them vulnerable to such phonological processes.

Systematically testing for this variable would be an instructive project. There are indications that phonological processes in some families (e.g., Australian) are less fusional than in others (e.g., Germanic or Indo-European in general). Still, given our present state of knowledge (or rather ignorance), we cannot confidently conclude that PERSON markers adjacent to NUMBER markers have remained separate because active fusional phonologies are lacking in precisely these languages or families.

12. If adjacency is not as conducive to fusion as could have been expected, and if phonologies cannot be invoked as convincingly accounting for the difference between fusing and non-fusing languages or families, the final inferential conclusion would be that phonological fusion simply is not the main, or at any rate not the only, force driving a scenario of morphological separation turning into cumulation. For bound PERSON and NUMBER, keener to cumulate than any other pair of categories, potent other mechanisms must exist effectuating cumulation.

13. One explanation is that this is to be attributed to the historical origins of phonologically or morphologically bound PERSON markers in independent pronouns, where PERSON would be already cumulated with NUMBER (ASSOCIATION, AUGMENTATION). This explanation is unsatisfactory (i) because independent PERSON forms themselves can be separative relative to NUMBER and vice versa (cf. Mandarin *wo-men* SPEAKER-ASSOC etc.; Turkish *b-iz* SPEAKER-ASSOC, *s-iz* ADDRESSEE-ASSOC, *on-lar* 3-PL etc.) and (ii) because the phonological and morphological processes attendant upon the grammaticalisation of such source forms are not

guaranteed to be structure-preserving.

And since independent PERSON forms are far more frequently cumulated with NUMBER marking than being separative, this inclination itself would call for an explanation.

14. We would suggest that morphological reanalysis, typically triggered by zero forms in paradigms, is an important mechanism of change from separation and cumulation, and sometimes the other way round. But for present purposes we must leave it at that: phonological fusion has traditionally been overestimated and morphological reanalysis been underestimated as instrumental in creating cumulative morphology.

15. Needless to add, this paper is a plea to terminologically distinguish between "cumulation" and "fusion" in morphological typology.

Abstract

For morphologically **separative** exponents to become **cumulative** by means of phonological **fusion** (with sandhi processes of various kinds obscuring and eventually obliterating internal morpheme boundaries), consistent or at any rate frequent **adjacency** is a precondition: the exponents of two co-occurring categories are not fused at a distance. Now, for the terms of any given pair of morphological categories, if we find their separative exponents equally often adjacent and non-adjacent across languages which have both categories, we may conclude, *ceteris paribus*, that phonological fusion cannot be such a forceful agent in transforming separation into cumulation – for adjacent exponents would otherwise be less frequent because, in the course of time, they would have become fused.

The pair of categories that we have investigated for present purposes are PERSON and NUMBER (or, more appropriately in connection with PERSON, ASSOCIATION or AUGMENTATION, formally sometimes distinct from and sometimes similar to plain nominal NUMBER).

For inflectional and clitical PERSON and ASSOCIATION markers it is far more common to be cumulated than to be separated, which distinguishes them from just about any other pair of morphological categories, insofar as separation is overall far more common than cumulation. The usual explanation is that this is to be attributed to the historical origins of phonologically or morphologically bound PERSON-and-ASSOCIATION markers in independent pronouns, with these categories already cumulated. This explanation is unsatisfactory (i) because independent PERSON forms themselves can well be separative relative to ASSOCIATION and vice versa (cf. Mandarin *wo-men* SPEAKER-ASSOC etc., Turkish *b-iz* SPEAKER-ASSOC, *s-iz* ADDRESSEE-ASSOC, *on-lar* 3-PL etc.) and (ii) because the phonological and morphological processes attendant upon the grammaticalisation of such source forms are not guaranteed to be structure-preserving. (And if independent PERSON forms are frequently cumulated with ASSOCIATION marking, this inclination itself would call for an explanation.)

At any rate, instances of separative bound marking of PERSON and ASSOCIATION – for all or some PERSONS, for all or some ASSOCIATION terms (augmented/plural, restricted- or unit-augmented/ dual, trial, paucal), for all or some series of bound forms (with verbs, nouns, adpositions, or other words as hosts/bases), for all or some contextual categories (tenses/aspects, with occasional splits between perfective and imperfective, or moods, with occasional splits between imperatives and indicatives) – are attested in a genealogically and areally diverse range of languages. There is a tendency for such separation, by comparison with more common cumulation, to run in families, which suggests some degree of diachronic stability.

One generalisation that emerges from our sample of currently some 100 languages is that separate PERSON tends to come before ASSOCIATION when both are bound, regardless of whether they are prefixes/proclitics or suffixes/enclitics or either. In the matter at issue, PERSON marking and separate ASSOCIATION marking are less frequently found to be adjacent than non-adjacent; but the difference is not really massive. From this line of reasoning, there is thus no overwhelming support for the assumption that morphological cumulation prototypically results from phonological fusion.

This negative result fits in with other evidence, not to be elaborated here, that fusion has traditionally been overestimated and morphological reanalysis (typically triggered by zero forms in paradigms) been underestimated as instrumental in creating cumulative morphology.

(Needless to add, this paper is a plea to terminologically distinguish between “cumulation” and “fusion” in morphological typology.)