

GEN.SG = NOM.PL: a mystery solved?

Pavel Caha
CASTL, Tromsø

September 2012

Abstract

This paper proposes an explanation for a homonymy between GEN.SG and NOM.PL, attested in a number of languages. The homonymy is interesting for theories of syncretism, because of the nature of the categories involved. Specifically, the two terms of syncretism do not form a natural class on any dimension (SG vs. PL and NOM vs. GEN). Working in the framework of Nanosyntax (Starke 2009, Caha 2009), I explain the syncretism by proposing that the formation of the plural involves a silent noun (GROUP), which requires a genitive case on its complement. I provide independent evidence for such a proposal, focussing on the fact that the syncretism is unavailable for pronouns and often also demonstratives.

1 Introduction

In Czech, there is a homophony between singular possessors and plural subjects.

- (1) a. plat učitelk-y
 salary teacher-FEM.GEN.SG
 ‘(female) teacher salary’
- b. Učitelk-y stávkují.
 teacher-FEM.NOM.PL. strike
 ‘(Female) teachers are on strike.’

The homophony holds for all feminine nouns in Czech, and a good part of neuters too. It is much rarer (though attested) in the masculine gender. I give a couple of examples below. (The choice of the sample is governed by the desire to illustrate the homophony for as many allomorphs as possible.)

- (2) Czech declension (partial)

	woman, SG.	woman, PL.	song, SG.	song, PL.	bone SG.	bone PL.	car SG	car PL
NOM	žen-a	žen-y	píseň-ø	písn-ě	kost-ø	kost-i	aut-o	aut-a
ACC	žen-u	žen-y	píseň-ø	písn-ě	kost-ø	kost-i	aut-o	aut-a
GEN	žen-y	žen-ø	písn-ě	písn-í	kost-i	kost-í	aut-a	aut-Ø
LOC	žen-ě	žen-ách	písn-i	písn-ích	kost-i	kost-ech	aut-u	aut-ech
DAT	žen-ě	žen-ám	písn-i	písn-ím	kost-i	kost-em	aut-u	aut-ům
INS	žen-ou	žen-ama	písn-í	písn-ěma	kost-í	kost-ma	aut-em	aut-ama

The fact that the syncretism occurs across various paradigms, and that it is replicated for various markers (-y, ě, -i, -a) makes it a good candidate for a pattern that is systematic, and reveals something important about the nature of the categories involved. The impression is strengthened by the observation that the same syncretism is replicated across various languages. Within Indo-European languages, one may find it for instance in Lithuanian, Romanian, Latin, Albanian or Old Irish – to give an indication of the spread (both geographical and temporal). The table below gives two illustrative paradigms from Lithuanian:

- (3) Lithuanian declension (partial) (Ambrasas et al. 2006, 116-7)

	cherry tree SG.	cherry tree PL.	bee SG.	bee PL.
NOM	vyšni-à	výšni-os	bìt-é	bìt-ès
ACC	výšni-a	výšni-as	bìt-ẹ	bìt-ès
GEN	výšni-os	výšni-ụ	bìt-ès	bìč-iụ

Even though the syncretism does not seem widely attested outside of Indo-European (cf. Wunderlich 2004), Skolt Saami (Ugro-Finnic) provides one of its most robust illustrations: the syncretism holds for all nouns in the language (two sample paradigms below).

(4) Skolt Saami declension (partial) (Feist 2010, 145,152)

	hole SG.	hole PL.	bee SG.	bee PL.
NOM	kää'pp	kää'v	puu'ttes	pottaz
ACC	kää'v	koo'vid	pottaz	pottsid
GEN	kää'v	koo'vi	pottaz	pottsi

The formal evidence that GEN.SG and NOM.PL form a class relevant for morphological marking contradicts the intuition that there are no other grounds for it to be so. Such an intuition is supported by several observations: first, there is little reason to think that in any of these languages, NOM-GEN syncretism has any significance within any given number. Second, there is also little reason to think that SG-PL has any significance within a particular case (in these languages). And lastly, the meaning of the two categories does not suggest any commonalities (though see Manzini and Savoia 2011).

Because of these three reasons, Baerman et al. (2002) include GEN.SG-NOM.PL syncretism into a category where “[f]ew would dispute that these patterns have come about by chance as a result of independent phonological developments, and [...] no Gesamtbedeutung should be sought.” Wunderlich (2004) says that GEN.SG=NOM.PL is “an accidental syncretism, caused by reasons other than underspecification[. It] can be captured by the meta-generalization that allows affixes to have the information <(+hr)N ∨ +pl>,” a simple disjunctive statement.¹ Yet other researchers suggest that it is exactly the “contradictory” meaning of the two categories what makes it possible for the syncretism to occur (the so called ‘polarity,’ see Béjar and Hall 1999, Lahne 2007, Wunderlich 2012).

An additional concern arises for approaches proposing that syncretism is restricted by contiguity.² The essence of such proposals is the idea that there is a geometrical arrangement of cells such that only contiguous regions may be targeted by syncretism. For such theories, the syncretism is problematic also for the reason that it is (apparently) non-contiguous.

Considerations presented in the preceding paragraphs could be summed up as follows: the homophony between GEN.SG and NOM.PL is systematic in a number of languages, but the two terms do not seem to have common meaning, and they are not contiguous. If this was true, the homophony would provide a good case in favor of a framework which imposes no (strict) restrictions on the terms of syncretism. For example, within the framework of Paradigm Morphology, Stump (2001) provides an easy way to state the facts using rules of referral; devices that link an exponent of one of the cell to the exponent found in another cell. In this particular case (using comparable Russian examples), Stump argues that such rules may go both ways, i.e., from GEN.SG to NOM.PL or the other way, yielding what he refers to as a bi-directional syncretism.

In this theoretical context, the current paper proposes an account of the syncretism that (i) provides a good reason why GEN.SG and NOM.PL form a natural class; and (ii) preserves a contiguity restriction on such a relation. This is achieved by the proposal that the formation of nominative plural involves a silent noun (GROUP) §2. In §3.1, I discuss a language (Mauritian Creole) which forms plurals by adding an overt noun (‘group’) to the singular form. In §3.2, I highlight a parallel between the distribution of the noun ‘group’ in Mauritian Creole (does not occur with pronouns) with the extension of the syncretism in Skolt Saam (does not extend to pronouns). Such a correlation provides an independent evidence for the proposal. In §4, I provide yet another type of evidence in favor of the proposal that plurals correspond to bi-nominal expressions. Specifically, I show that the puzzling agreement system in Bayso provides evidence to this effect.

2 Towards a solution

Let me start by de-trivializing the observation that GEN.SG=NOM.PL syncretism is problematic for theories that restrict syncretism by contiguity. Specifically, the question arises why the cells should be ordered as shown in (5), rather than as shown in (6).

(5) Classical arrangement

	SG.	PL.
NOM	A	C
ACC	B	E
GEN	C	F
LOC	D	G

(6) New arrangement

¹(+hr)N roughly means that the morpheme may express subject-like elements in nominal environment.

²See, e.g., McCreight and Chvany 1991, Plank 1991, Johnston 1996, Trosterud 2004, Wiese 2003, Bobaljik 2007/2012, Caha (2009), Starke 2009, Pantcheva 2010, 2011, Taraldsen 2010, Vangsnes in press for relevant discussion).

	SG.	NOM.PL	ACC.PL	GEN.PL
NOM	A	–	–	–
ACC	B	–	–	–
GEN	C	C	E	F
LOC	D	–	–	–

Recall that apart from $\text{GEN.SG}=\text{NOM.PL}$, there are otherwise very few (if any) syncretism relations between sg and pl. Hence, the ‘new’ arrangement loses nothing (or very little) in terms of syncretism coverage. However, the new arrangement seems close to becoming vacuous. Why? The answer is that apart from syncretism alone, the linear arrangement of cells is expected to express natural classes. For example, ACC.SG and ACC.PL obviously form a natural class (they are ACC). However, the paradigm space depicted in (6) does not capture this. The suspicion is that once we are allowed to draw paradigm spaces ‘as needed,’ the contiguity hypothesis becomes an exercise in geometry, and loses interest and relevance to grammatical theory.

In frameworks which understand contiguity in terms of feature decomposition, the problem becomes even more apparent. To see that, I turn to the implementation of contiguity in the framework I will be assuming here, Nanosyntax (see Starke 2009, 2011). Let me mention in advance that the specific nature of the problem is going to be instructive, and it directly leads to the solution I propose in this paper. Consider first some background.

2.1 Syncretism and contiguity in Nanosyntax

The particular way Nanosyntax encodes contiguity is in terms of ‘cumulative’ feature decomposition (see Caha 2009, to appear). In particular, if there is a linear arrangement (say $\text{NOM} - \text{ACC} - \text{GEN}$) where only contiguous regions exhibit syncretism (see (7-a)), the framework encodes this by the decomposition shown in (7-b-d):

- (7) a. linear sequence: $\text{NOM} - \text{ACC} - \text{GEN}$
b. $\text{NOM} = [\text{X}]$
c. $\text{ACC} = [\text{X}, \text{Y}]$
d. $\text{GEN} = [\text{X}, \text{Y}, \text{Z}]$

With such a decomposition, it is quite easy to see that there is no way to state a syncretism between $[\text{X}, \text{Y}, \text{Z}]$ and $[\text{X}]$ to the exclusion of $[\text{X}, \text{Y}]$. This is because $[\text{X}, \text{Y}]$ is ‘more similar’ to $[\text{X}]$ than $[\text{X}, \text{Y}, \text{Z}]$; hence, in case $[\text{X}, \text{Y}, \text{Z}]$ and $[\text{X}, \text{Y}]$ are distinct, $[\text{X}]$ is expected to pattern with $[\text{X}, \text{Y}]$, rather than $[\text{X}]$.

With the basic idea in place, let me give some more details about how insertion works in Nanosyntax, as this will become relevant later. In particular, Nanosyntax is a realizational theory of morphology, which means that lexical entries are constructed as rules which take the syntactic structure as their input, and produce phonological representations as their output. Whether an entry applies in a given case is determined by the so called Superset Principle:

- (8) *The Superset Principle (Starke 2009):*
A lexical entry is inserted into a node if it contains that node.

That means that the entry for a genitive marker is as given in (9):

- (9) $\text{GEN} \Leftrightarrow [\text{X}, \text{Y}, \text{Z}]$

Note, however, that the ‘genitive’ marker may also be inserted in $\text{acc} [\text{X}, \text{Y}]$ and $\text{nom} [\text{X}]$, due to The Superset Principle. That is because the rule allows insertion in all cases which are contained in the entry (9). The consequence is that an entry such as (9) may be inserted in GEN , ACC and NOM .

Whether it actually surfaces in those cases depends on whether there are any competitors, and what the competitors are. For example, suppose there is the additional entry (10):

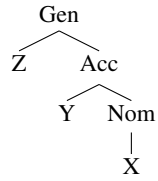
- (10) $\text{NOM} \Leftrightarrow [\text{X}]$

This entry may not apply in GEN and ACC (it does not contain those features). Therefore, these cases are spelled out using the rule (9). However, in the NOM , both the entry (9) and (10) may apply. A competition arises with the consequence that the more specific entry (i.e., (10)) wins, and it is chosen as the ultimate spell out of NOM .

In addition to cumulative decomposition, Nanosyntax proposes that each of the features is a separate head in the syntactic tree, with the heads ordered along a hierarchy called the functional sequence. So in fact, the representation (10-d) looks like (11). The series of the non-terminal projections (the functional sequence) corresponds to the geometric arrangement of purely spatial accounts.³

³Note that Nanosyntax adopts the phrasal spell-out hypothesis, according to which lexical entries may spell out whole phrases. The complex syntactic layering thus does not entail complex morphology (at least not any more than the very idea of cumulative decomposition).

(11) GEN =



Summing up the discussion, the following two statements characterize the approach to syncretism adopted in Nanosyntax:

- (12) Syncretism in Nanosyntax: two rules of thumb
- If $C1=C2$, then either $f(C1)$ contain $f(C2)$, or the other way round.
 - Each feature is a head in the tree

The decision to model syncretism as a specific instance of structural containment will become relevant, because such representations allow for an independent verification by morphological containment relations. For example, in many languages, *INS* is coded the same as *COM* (Stolz et al. 2006). In Nanosyntax (see (12-a,b)), this means that either (13-a) or (13-b) holds:

- (13)
- $COM = [X [INS]]$
 - $INS = [X [COM]]$

It turns out that if there is a morphological containment relation between *INS* and *COM*, then *COM* always includes *INS* as its component part (Stolz et al. 2009, 607). From this, we may conclude that (13-a) is correct. Repeating the crucial message: Nanosyntax analyzes syncretism as a specific manifestation of structural containment.

2.2 GEN.SG=NOM.PL in Nanosyntax

Applying the general guidelines (12-a,b) to the case we are interested in here (*GEN.SG=NOM.PL*), the representation we arrive at is shown in (14):



There is one thing right about this, and (at least) one thing wrong. The right thing is that there are languages (North Saami, Estonian) where such a containment relation is apparent in the morphological make-up. I illustrate this in (15) for North Saami. This observation then falls in line with the general observation that syncretism and containment go hand in hand for a given relation (but in different languages).

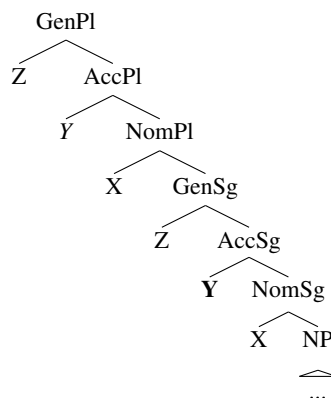
(15) North Saami declension (Nickel 1990)

	house, SG	house, PL	reindeer, SG.	reindeer, PL
NOM	viessu	viesu-t	boazu	bohcco-t
ACC	viesu	viesu-id	bohcco	bohccu-id ⁴
GEN	viesu	viesu-id	bohcco	bohccu-id
ILL	vissui	viesu-ide	bohccui	bohccu-ida
LOC	viesus	viesu-in	bohccos	bohccu-in
COM	viesuin	viesu-iguin	bohccuin	bohccu-iguin
ESS	viessun	—	boazun	—

The ‘wrong’ thing about the proposal in (14) is that it leads to a problem with recursion. I show this in (16). The reasoning leading to the tree is this: the feature by which *ACC.SG* is ‘derived from’ *NOM.SG* (the boldfaced *Y* in (16)) has to be the same as the feature by which *ACC.PL* is derived from *NOM.PL* (the *Y* in italics). And similarly for other cases; hence the picture shown in (16), where there is a recursion of all the three case features.

⁴According to (Nickel, 1990, 78), the form *bohccu* is due to an orthographic convention such that “*o+i* is written *ui*.”

(16)



However, (16) is not a legitimate sequence of categories: it has been independently established in syntactic research that the functional sequence of projections is an *irreflexive* ordering (A may never dominate A; see Cinque 1999, Starke 2004). This principle of projection building is violated in (16) (where the projection of Y dominates that of Y). Hence, it seems that an attempt to encode this syncretism in Nanosyntax leads to a contradiction with its own assumptions; as Manzini and Savoia 2011, 115 observe, “the syncretism of [...] nominative plural and genitive singular [...] constitutes a problem for the Nanosyntax model, which by construction is incapable of capturing it.”

Stepping outside of Nanosyntax, however, does not help either. Suppose, for instance, that in order to avoid a clash between the representation (16) and our independent findings about the nature of the functional sequence, we give up the idea that features project in a syntactic tree. But then the representation of ACC.PL will have to be a set that contains two instances of the feature Y. But there is no such object in set theory: the set {Y, Y} is identical to the set {Y}.

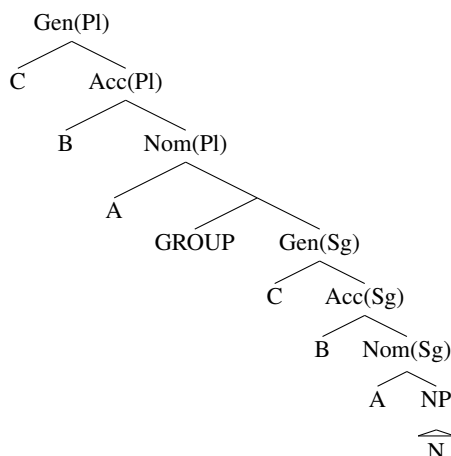
In general, it seems that any feature based framework will have to construct NOM.PL by minimally modifying the feature specification of GEN.SG. However, GEN.PL will always have to be related to NOM.PL in the same way as GEN.SG to NOM.SG. Hence, the ‘recursion’ problem arises independently: the same relation (between GEN and NOM) will be present twice in the same feature structure.

The discussion leaves us with two relevant conclusions. First, it seems that the idea of containment between GEN.SG and NOM.PL is on the right track (because of North Saami and its kin). What we need to do is to remove the recursion problem.

2.3 There is recursion ... well ... because there is recursion

The following statement provides a solution to the puzzle: Plurals in Czech (and elsewhere) are bi-nominal structures, where the genitive singular is an argument of a silent quantity noun, represented in the tree below by the item GROUP. The presence of a second NOUN is crucial; the second noun provides the only explanation for recursion.

(17)



While solving one problem, it may seem that a number of other problems get created. All the problems stem from the straightforward expectation that the whole structure will behave like a quantity phrases of the sort ‘majority of voters,’ except that the head noun is silent. This expectation has two empirically observable consequences. (i) The dependent noun is going to be in the genitive no matter the syntactic environment, because it is the silent head noun which bears the various case affixes appropriate for the role of the whole phrase. (ii) All nouns will be subject to GEN.SG=NOM.PL syncretism, because they are all possessors of the silent head.

Both of these expectations are wrong. As highlighted at the start, only a subset of Czech paradigms shows the GEN.SG=NOM.PL homophony. I give an example of one such paradigm below in (18) (see ‘castl’), side-by-side with a ‘well-behaved’ paradigm (‘machine’). The plural paradigms also illustrates the incorrectness of the expectation that the plural noun is not going to change depending on its syntactic environment: no matter whether the NOM.PL is identical to GEN.SG or not, the noun obviously reflects features of its larger syntactic environment (as opposed to expressing only the possessive relationship).

(18) Czech masc decl

	machine, SG.	machine, PL.		castl SG	castl PL
NOM	stroj-ø	stroj-e		hrad-ø	hrad-y
ACC	stroj-ø	stroj-e		hrad-ø	hrad-y
GEN	stroj-e	stroj-ů		hrad-u	hrad-ů
LOC	stroj-i	stroj-ích		hrad-u	hrad-ech

2.4 Fine tuning the proposal

Both problems disappear if we weaken the expectation that the structure with a silent noun GROUP in it will behave exactly like ordinary possessive structures. Specifically, I propose that the noun GROUP gives rise to a specific ‘agreeing possessor’ construction, the so called Suffixaufnahme (Plank 1995). Common nouns, on the other hand, have no possessor agreement. The fact that different possessive structures may be expressed differently within a language should not come as a big surprise, because such variation is independently attested.

To see what Suffixaufnahme looks like, consider first the data in (19). As the contrast between (19-a) (*marun-ngan-i*) and (19-b) (*marun-ngan-ku*) shows, the genitive marking of the possessor is followed by an agreement marker that tracks the case of the head noun. The general pattern is shown in (19-c).

- (19) Lardil, (Richards 2007)
- a. [*marun-ngan*]-i wangalk-i
 boy-GEN-ACC boomerang-ACC
 ‘the boy’s boomerang’
- b. [*marun-ngan*]-ku maarn-ku
 boy-GEN-INS spear-INS
 ‘with the boy’s spear’
- c. [[N-*gen*]-agr]

Another relevant observation is that in some languages, the genitive features and the agreement features are expressed by a single portmanteau. The examples below are from Bezhta (Caucasian), taken from Plank (1995, 71). In (20-a), -s is the genitive ending for possessors whose head is in the nominative case. When the head noun is in an oblique case (DAT in (20-b)), the genitive marker is -la.

- (20) a. [abo-s] is
 father-GEN1 brother
 ‘father’s brother’
- b. [abo-la] is-t’i-l
 father-GEN2 brother-OBL-DAT
 ‘to father’s brother’

Putting these two pieces together, we obtain the relevant syntax that is manifested in the Czech NOM.PL (and analogous examples). Specifically, I propose that the possessor of the silent noun GROUP agrees with the head. The syntactic tree reflecting the proposal is given in (21). In the tree, the genitive comes adorned with an additional constituent bearing agreement features (boldfaced).



The proposal in (21) solves both of the problems we have started from. The reason why the NOM.PL (i.e., a type of a possessor) does not stay invariantly genitive throughout the paradigm is because of the presence of Agr, which tracks the case of the head, and represents it on the possessor.

Further, GEN.SG and NOM.PL are not exactly alike. (The formula is this: NOM.PL=GEN.SG+AGR.) This explains the cases where NOM.PL is distinct from GEN.SG. To show that in greater detail, consider the lexical specification of a marker that acts as a nominative plural. In particular, the marker spells out both NOM.AGR and the embedded genitive marking, as shown in (22):

(22) $\text{NOM.PL-Marker} \leftrightarrow [\text{Agr} [\text{gen}]]$

As highlighted above, in Nanosyntax, any entry may spell out structures that correspond to their full specification, or structures that are a subset of the lexical specification (recall The Superset Principle (8)). The $\text{NOM.PL}=\text{GEN.SG}$ homophony then arises when an entry such as (22) is allowed to make use of this possibility, and spell out a structure that only corresponds to the GEN part. This happens in paradigms in which there is no ‘better’ competitor for the embedded structure. In case there is, this competitor wins, and the NOM.PL entry is restricted to plural only.

2.5 Summing up

At this point, we are done accounting for the $\text{GEN.SG}=\text{NOM.PL}$ homophony. The proposal says that the homophony arises as a consequence of the fact that the genitive case reflects the syntactic role of the noun: it is a possessor of a silent noun GROUP . In addition, the possessor agrees in case with the head noun, and this agreement is represented on the possessor in a separate constituent, generated on its top. This proposal explains why the marking of the possessor does not stay constant regardless of the case assigned to the head; in this regard, the construction is analogous to the situation found in Bezhta (see (20)) and other languages. At the same time, we get a handle on how to handle cases where GEN.SG and NOM.PL are distinct.

Crucially, the theory achieves this without giving up on contiguity. Instead, the traditional hypothesis about the organization of the paradigm space (SG and PL are two parallel columns) is rejected. NOM.PL is built directly on top of the genitive singular, and these two terms thus end up ‘contiguous’ in the technical sense of the term.

3 Where $\text{GEN.SG}=\text{NOM.PL}$ systematically fails

The remainder of this paper is dedicated to showing that this proposal directly accounts for additional facts, not captured by the alternatives. In this section, I show that in a number of languages, the $\text{textsc{gen.sg}=\text{nom.pl}}$ homophony fails to apply to pronouns and demonstrative determiners. I show how this fact follows from the hypothesis developed here, namely that NOM.PL involves a possessive structure with a silent head GROUP . Systematic restrictions like these, however, come as a surprise in approaches where the identity between GEN.SG and NOM.PL can be stipulated by a simple rule (Stump 2001) or a disjunctive statement (Wunderlich 2004). The discussion is thus intended to show that the current theory is not only more restrictive (preserves contiguity), it is also empirically superior to its alternatives.

I start by introducing a set of facts from Mauritian Creole, which form the basis for further discussion.

3.1 Mauritian Creole *bann*

In Mauritian Creole, there is a morpheme (*bann*) which has two major uses. In its first use, it means something like a ‘group, community.’ I show this below:

- (23) Dan nou **bann**, maryaz pa fer sa kalitela, dan zot **bann**, ki fer koumsa
 in our community marriage not do that way, in their community, COMP do that way
 ‘In our community, marriages aren’t done that way, it’s in their community that they are done that way’
 (Guillemin 2011, 186)

In its second use, it works as a plural marker. The examples (24-a-c) illustrate its use in this capacity for subjects, objects and prepositional objects respectively.

- (24) Mauritian Creole (Guillemin 2011, 190-1)
- a. **Bann** frizider vid.
 PL fridge empty
 ‘The fridges are empty.’
 - b. kifer li pa kapav donn lord lalinn, **bann** planet, **bann** zetwall, ...?
 why she NEG can give order moon PL planet, PL stars
 ‘Why can’t she give orders to the moon, the planets, the stars ... ?’
 - c. Enn kantite sa kalite prezize lor **bann** langaz finn demanti ...
 a lot DEM type prejudice on PL language ASP denounce ...
 ‘Many biased views on language have been exposed ...’

This polyfunctionality is interesting for several reasons. First, it shows that it is possible for a ‘group’ type noun to serve as a plural marker, an essential part of the current proposal. Two additional facts concerning *bann* and its syntax become relevant shortly. The first fact is that the demonstrative determiner *sa* occurs to the left of *bann*:

- (25) a. sa zom
 this man
 ‘this man’

- b. sa bann zom la
 this PL man SP
 ‘these men’ (Guillemin 2011, 256)

The second fact is that the plural of pronouns does not have *bann*. The following examples show that.

(26) Mauritian Creole pronouns (Adone 1994, 36)

- a. 1: sg. *mo* — pl. *nu*
 b. 2: sg. *to* — pl. *u*
 c. 3: sg. *li* — pl. *zot*

This is quite likely related to the contrast between nouns and pronouns when it comes to the meaning of plural. While ‘houses’ can be considered a group of individuals of which the predicate ‘house’ holds, the plural forms of pronouns cannot be construed that way. (*We* is not a group of individuals of which the predicate ‘me’ holds.) Regardless of whether this explanation is right or wrong, the absence of *bann* with pronouns is all that matters for the following discussion.

3.2 Where GENSG=NOMPL fails in Skolt Saami

As highlighted above (see (4)), Skolt Saami has GENSG=NOMPL for all nouns. But there are two sets of expressions in the language where the syncretism does not occur. The first domain is in the realm of pronouns.

(27) Skolt Saami pronominal declension (partial) (Feist 2010, 251-2)

	1 SG	1 PL		3 SG	3 PL		refl SG	refl PL
NOM	mon	mij		son	sij		jiōčč	jiijj
ACC	muu	mi’jjid		suu	si’jjid		jiijâs	jiijjâz
GEN	muu	mij		suu	sij		jiijâs	jiijjâz

This fact follows from the current account, coupled with the observation that pronominal plurals lack GROUP, as indicated by the Mauritian Creole data, see (26)). That is because GEN.SG=NOM.PL is actually caused by the presence of the silent noun GROUP; with the noun missing, the syncretism disappears.

The second domain where GEN.SG=NOM.PL fails to occur in Skolt Saami is in the domain of demonstrative determiners.

(28) Skolt Saami demonstrative determiners (partial) (Feist 2010, 251-2)

	PROX, SG	PROX, PL		DIST, SG	DIST, PL
NOM	tât	tâk		tut	tuk
ACC	tân	tîd		tun	tuid
GEN	tân	tâi		tun	tui

The explanation emanates from the syntactic position of demonstratives in such structures, shown in (29) (recall (25)):

(29) [**DEM** [**GROUP** [of α]]]

In particular, I assume that in this structure, an agreement relation is established with the closest noun: the silent GROUP. If that is so, the demonstrative reflects the features of the noun GROUP, rather than the features of the complement. (Agreement indicated by boldface.) As a consequence, the embedded genitive is not reflected on the demonstrative, and the syncretism fails.

3.3 Romanian feminine declension

Similar observations hold for Romanian. The table below shows the declension of the Romanian phrase ‘a talented actress.’ The shading indicates homophony between GEN.SG. and NOM.PL.

(30) Romanian (Cojocaru nd)

	INDEF	actress	talented
NOM/ACC.SG.	o	actriț-ă	talentat-ă
GEN/DAT.SG.	unei	actriț- e	talentat- e
NOM/ACC.PL.	niște	actriț- e	talentat- e

Interestingly, the syncretism pertains only to the inflection of the noun and the adjective (boldfaced). The article is distinct in the two cases. Assuming a picture like that in (32) gives us a clear idea why this is so: the adjective

and the noun are in the scope of the plural quantificational noun, and they are assigned genitive singular. The article is outside of the scope, and it is not subject to the assignment, or the homophony:

- (31) [niște [GROUP [actriț-e talentat-e]]]
INDEF actors talented
a GROUP of talented actresses

The failure of GEN.SG=NOM.PL on demonstrative determiners and articles presents evidence for the presence of a silent noun, because in such languages, the scope of the GEN.SG=NOM.PL homophony coincides with the scope of the silent noun. But there is a twist: there are languages where demonstratives do exhibit the GEN.SG=NOM.PL homophony. How is this possible?

3.4 Omnivorous demonstrative agreement

In order to understand that, let me start from the observation that descriptively, there are at least two types of agreement. First, there are languages where agreement consistently tracks the features of a single controller (say, the subject), no matter what its features are. An example of such a system is the participle agreement in Czech, with an example given in (32). In (32-a), both the subject and the object are singular, as well as the participle. In (32-b), the object is plural, but the participle stays singular, reflecting faithfully the features of the subject. The participle agreement becomes plural only when the subject is plural, as in (32-c).

- (32) a. Petr-ø pomaloval-ø zed'-ø.
Petr-SG painted-SG wall-SG
 b. Petr-ø pomaloval-ø dvě zd-i.
Petr-SG painted-SG tw wall-PL
 c. Petr-ø a Karel-ø pomaloval-i zed'-ø.
Petr-SG and Karel-SG painted-PL wall-SG

However, there are also languages where agreement may reflect various controllers, depending on their markedness. An example of such a system is given in (33), and comes from Abruzzese (D'Alessandro and Roberts 2010). As a base line, the example (33-a) shows the singular form of the past participle. Interestingly, when the object is plural, the participle agreement changes, and becomes plural. However, this is not due to the fact that the participle agreement tracks (only) objects. This can be seen in (33-c). Here, the object is singular, but the participle is plural all the same. This is caused by the plurality of the subject. Therefore, the description is that the participle agreement is plural whenever either subject or object (or both) are plural.

- (33) a. Giuwanne a pittate nu mure
John has painted.sg a wall
 b. Giuwanne a pittite ddu mure
John has painted.PL two walls
 c. Giuwanne e Mmarije a pittite nu mure
John and Mary have painted.PL a wall

A plausible explanation of such patterns relies on the notion of markedness or specificity. The idea is that plural is more marked than singular; it has more features. Agreement in Abruzzese is sensitive to this information, and tracks the more specific item. Nevins (2011) calls such a pattern 'omnivorous' agreement.

There is evidence suggesting that the same variation (fixed controller vs. omnivorous) obtains for demonstratives. To see that, consider the examples in (34). They come from Czech (34-a) and Polish (34-b), two closely related languages, and show how the two languages render the phrase 'those five books,' when it occurs in the object position.

- (34) Czech vs. Polish demonstrative agreement
 a. t-ěch pět-ø knížek
those-GEN.PL five-ACC books.GEN.PL
'those five books'
 b. t-e pięć-ø książek
these-ACC.PL five-ACC books.GEN.PL
'these five books'

What is common between the examples is that the numeral bears the case assigned to the whole phrase (acc), and requires the genitive case on its complement. Ionin and Matushansky (2006) argue that this reflects the fact that Slavic numerals like 'five' are in fact nouns. Crucially, the examples differ in the case represented on the demonstrative. While the Polish demonstrative is in the accusative and reflects the features of the numeral, the Czech demonstrative is in the genitive case, reflecting the features of the embedded noun.

This can be explained under the hypothesis that the Polish demonstrative tracks the features of a dedicated controller, the head noun ‘five.’ On the other hand, the Czech demonstrative exhibits omnivorous agreement, and reflects the features of the more marked controller (the GEN), and apparently skips the accusative numeral.

Given the behavior of the Czech demonstrative, we in fact expect that in some languages, demonstratives will also be able to skip the silent GROUP, and reflect the features of the embedded genitive. Lithuanian is a case in point, with the relevant data in (35). The first row of the table shows a nominative singular of the phrase ‘that wooden hand.’⁵ In the second row, we see the genitive singular form (shaded). It is fully identical to the nominative plural row, which can be found below.

(35) Lithuanian (Press 2004, 14-8)

	that sg.	wooden, sg.	hand sg.
NOM.SG	tà	medinė	rankà
GEN.SG	tõs	medinẽs	raĩkos
NOM.PL	tõs	medinẽs	raĩkos
GEN.PL	tũ	medinių	raĩkų

In sum, the point of this section has been to show that there are two types of agreement: agreement that always tracks the features of a dedicated controller, and agreement which may track multiple controllers. In the latter case, the features expressed on the target depend on markedness: the agreement features reflect the more marked controller.

Such a variation is replicated in the domain under discussion. While in some languages, demonstratives and determiners track the features of the head noun GROUP (Romanian, Skolt Saami), this is not the only possibility. The demonstrative may as well skip the silent noun, and agree with the embedded genitive (Lithuanian). As a result, this gives us an implicational hierarchy of likelihood where GEN.SG=NOM.PL OCCURS:

(36) The likelihood of GEN.SG=NOM.PL: N > Dem

In words, if demonstratives show GEN.SG=NOM.PL syncretism, then nouns do as well. But it may be the case that demonstratives fail to have such a syncretism where nouns have it; this is because in languages where demonstrative agreement is not omnivorous, it fails to reflect the embedded genitive.

3.5 Summary

In the current section, I introduced the Mauritian Creole plural morpheme *bann*, homophonous with the noun ‘group.’ Interestingly, this morpheme is absent with pronouns. I suggested that its absence is the reason for the corresponding lack of homophony between GEN.SG and NOM.PL forms of Skolt Saami pronouns, a language where the homophony otherwise targets all nouns.

Another domain where the GEN.SG=NOM.PL homophony frequently disappears is on demonstrative determiners. This receives an explanation from the position of demonstratives in such structures, i.e., higher than the (silent) plural noun GROUP. In this high position, the demonstrative may agree with the (silent) noun, and fails to reflect the genitive case of its complement. Therefore, no homophony arises.

In the next section, I follow this line of reasoning further, and I argue for the bi-nominal structure of plurals from the observable effects the GROUP noun may have on agreement.

4 How agreement reveals a bi-nominal recursive structure

The discussion is going to revolve around Bayso, a Cushitic language with an intriguing agreement system (Corbett and Hayward 1987). I argue that the peculiarities of the system may be understood under the hypothesis that its various plural markers are nouns with an inherent gender.

I think that the best way to understand the Bayso system is to look first at its pronouns; these are shown in the table below. The terms individual reference and multiple reference are self-explanatory; I follow Corbett and Hayward (1987) and use them instead of the common labels singular and plural for reasons that will become clear shortly.

(37) Bayso pronouns (Corbett and Hayward 1987, 12)

	individual reference	multiple reference
MASC	úsu	íso
FEM	ése	íso

⁵I could not find declined complex examples in the grammars, so the phrases arise by juxtaposition of the forms as found independently in the paradigms.

The pronouns *úsu*, *ése* and *íso* trigger each a particular agreement form of the verb; let me call these MASC, FEM and PLURAL.

The nouns in the language have a basic form (a form without affixes). This form can refer to a single individual, as well as a whole group. The absolute majority of these basic (number-less) forms falls into two classes: masculine or feminine. I illustrate the two classes below.⁶

- (38) Nominal agreement (basic form)
- a. *lúban hudure*
lion slept.MASC
'lion(s) slept.'
 - b. *kimbír hudurte*
bird slept.FEM
'bird(s) slept.' (Corbett 2000, 181-2)

To create forms with singular reference, the suffix *-(ti)ti* is used. The suffix does not change agreement class membership.

- (39) Nominal agreement (individual reference)
- a. *lubán-titi hudure*
lion slept.MASC
'A lion slept.'
 - b. *kimbír-titi hudurte*
bird slept.FEM
'A bird slept.' (Corbett 2000, 181-2)

The surprising thing is that in the multiple reference form, both genders trigger MASC agreement.⁷

- (40) Nominal agreement (multiple reference)
- a. *luban-jool hudure*
lion-MULTI slept.MASC
'The lions slept'
 - b. *kimbir-jool hudure*
bird-MULTI slept.MASC
'The birds slept.' (Corbett 2000, 181-2)

We know that the predicate is in the MASC form, because the plural agreement (as revealed by the multiple reference pronoun) would be different (*hudureene*). In order to encode these facts, we have to somehow (i) specify each root for the gender it has in its base form (ii) specify the multiple reference affix *jool* as masculine. Doing so reveals a 'recursion' problem. In other words, in the multiple reference forms, gender is represented twice:

- (41)
- a. [[*luban.MASC*] *jool.MASC*]
 - b. [[*kimbir.FEM*] *jool.MASC*]

The proposal that plurals in Bayso are bi-nominal structures explains the issue; once again, we get the effects of recursion because there actually is recursion. At the same time, assuming that *jool* (the overt counterpart of GROUP) is a noun which triggers masculine agreement explains the apparent oddity of the fact that multiple reference forms do not trigger plural agreement. Similarly, assuming once again that pronouns do not have the noun GROUP explains the differential behavior of pronominal vs. nominal multiple reference forms.

The complexity of the system does not stop here. When it comes to the formation of the multiple reference form, there are two additional minor classes of nouns. One of them consists of nouns whose multiple reference forms actually do trigger plural agreement; I give them below.

⁶There are several exceptional nouns that trigger plural agreement in their basic forms; see Corbett and Hayward 1987 for a discussion. These are mainly body part names, i.e., names of objects that characteristically come in more than a single exemplar.

⁷This is also the reason I avoid using the term plural for the multiple reference forms; it sounds rather awkward to say that plural nouns do not trigger plural agreement.

(42) Nouns with plural multiple reference agreement (Corbett and Hayward 1987, 13)

base form	base agr	multiple reference form	meaning
baal	M	baal- <i>allo</i>	‘feather, leave’
suul	M	suul- <i>allo</i>	‘nail, claw’
fer	M	fer- <i>erroo</i>	‘finger, toe’
gilib	M	gilb- <i>oo</i>	‘knee’
nébe	F	nebe- <i>bboo</i>	‘knee’
aayo	F	aayo- <i>os</i>	‘mother’
ilk- <i>oo</i>	PL	ilk- <i>oo-l</i>	‘teeth’

As Corbett and Hayward (1987) point out, apart from their exceptionality in agreement, these nouns lack the regular plural suffix, and mostly end in *-o*: “The irregularity of the nouns in the table [(42)] is one of number: they must be marked as taking plural agreement when they are in the multiple reference form. But these nouns must in any case be marked as irregular. The first six take an irregular suffix [...], the last one [...] takes irregular agreement with the singular reference form” Corbett and Hayward (1987, 13).

Thus, for the first six nouns, we observe a correlation between the lack of masculine agreement and the lack of the regular suffix. This supports the hypothesis that the regular suffix *jool* is indeed the source of the masculine agreement, and by triggering such an agreement, falls in the same class as nouns.

The proper analysis of *ilk-oo* ‘teeth’ is somewhat unclear. Corbett and Hayward (1987) seem to assume that *-l* is an allomorph of the regular suffix *-jool*. There are two options how this can be interpreted. Either *-l* is simply a different noun in the same function, in which case the correlation between the lack of *-jool* and the lack of MASC agreement still holds. Alternatively, the occurrence of *-l* is phonologically determined, in which case we could expect MASC agreement. The fact that it does not occur would need a special explanation; a plausible line of analysis would rely on the presence of *-oo* in the singular reference form, which seems to correlate with the fact that the singular reference form exceptionally triggers PL agreement.

Another group of exceptions is presented by nouns which take feminine agreement in their multiple reference form.

(43) Nouns with feminine multiple reference agreement (Corbett and Hayward 1987, 13)

base form	base agr	multiple reference form	meaning
aar	M	aar- <i>aar</i>	‘ox, bull’
jarsa	M	jars- <i>olee</i>	‘old man’
abbi	M	abbi- <i>laal</i>	‘brother’
abba	F	abba- <i>laal</i>	‘sister’

These nouns, once again, lack the regular suffix *-jool* (or *-l*). As Corbett and Hayward (1987, 13-4) observe, “the last two take the suffix *-laal* and are the only ones to do so.” The other two nouns may be analysed as nouns with an irregular collective suffix. Thus, once again, there is a strong connection between the suffix involved, and the agreement form taken.

These facts taken together provide interesting evidence for the nominal nature of the plural affixes. It seems that the agreement system is not sensitive to the meaning of the forms, but rather to the particular affix that is used in order to derive multiple reference forms. The way these affixes govern gender agreement is arbitrary, and therefore, reminiscent of nouns, whose gender specification is known to be partly arbitrary in the same sense. This in turn provides evidence for the nominal status of the plural affixes. In general, it seems that the bi-nominal structure I offer here has effects that reach beyond the GEN.SG=NOM.PL homophony, providing independent support for the original idea.

5 Summary and conclusions

The paper started from the observation that GEN.SG=NOM.PL raises non-trivial challenges for theories of syncretism. I then went on to show how syncretism is treated in Nanosyntax, and showed that when we attempt to account for this syncretism in a purely mechanical way, a recursion problem arises. I took this result at face value, and proposed that plurals in the languages with the syncretism and elsewhere are bi-nominal recursive structures. The solution to the recursion problem is thus simple: we get effects of recursion because there in fact is a real recursion. Later on, I have showed that the recursion problem arises quite independently in Bayso, and the same solution has been applied.

The beneficial consequences of the solution have been explored in the paper. First, the solution allows us to account for the syncretism in a theory which is restrictive in the sense that it derives the contiguity restriction on syncretism (non-contiguous syncretisms cannot be represented in such a theory). Second, I argued that the GEN.SG=NOM.PL syncretism is only one out of many possible manifestations of a bi-nominal structure.

First, the proposal directly accounts for languages where NOM.PL is morphologically based on GEN.SG (North

Saami, Estonian). Second, the bi-nominal structure is obvious in languages where plurals are formed using a morpheme homophonous to a noun ‘group’ (Mauritian Creole). Third, building on the observation that Mauritian Creole pronouns lack the ‘group’ morpheme, an explanation is provided for the fact that GEN.SG=NOM.PL fails to arise for pronouns (Skolt Saami and elsewhere). Fourth, the fact that the homophony often disappears on demonstrative determiners is accounted for by the proposal that demonstratives may agree with the ‘group’ noun. Finally, the proposal allows us to understand apparently ‘crazy’ agreement systems where each plural (i.e., multiple reference) belongs to an agreement class of its own.

If correct, the account allows us to draw some general conclusions concerning contiguity. In particular, it seems that evaluating the correctness of the hypothesis relies on a pre-established grammatical space. Traditional conceptions of this space may be too simplistic (singular and plural may not be parallel columns); I have argued that the structure of plural is more complicated than standardly believed, with the consequence that the representation of GEN.SG forms a direct input for NOM.PL. Last but not least, if it is correct that pronouns may lack the noun GROUP, then the grammatical space against which contiguity is evaluated may be different for nouns and pronouns. That is because for pronouns, GEN.SG does not enter the formation of NOM.PL.

References

- Adone, Dany. 1994. *The acquisition of Mauritian creole*. John Benjamins, Amsterdam.
- Ambrasas, Vitautas, Emma Geniušienė, Aleksas Girdenis, Nijolė Sližienė, Adelė Valeckienė, Elena Valiulytė, and Dalija Tekorienė. 2006. *Lithuanian grammar*. Baltos lankos, Vilnius.
- Baerman, Matthew, Dunstan Brown, and Greville G. Corbett. 2002. Case syncretism in and out of Indo-European. In *CLS 37: The Main Session. Papers from the 37th Meeting of the Chicago Linguistic Society, Vol. 1.*, edited by M. Andronis, C. Ball, H. Elston, and S. Neuvel, p. 1528. Chicago Linguistic Society, Chicago.
- Béjar, Susana and Daniel Currué Hall. 1999. Marking markedness: The underlying order of diagonal syncretisms. In *Proceedings of the 1999 Eastern States Conference on Linguistics*, edited by R. Daly and A. Riehl. Cornell Linguistic Circle Publications, Ithaca, N.Y.
- Bobaljik, Jonathan. 2007. On comparative suppletion. Ms., University of Connecticut.
- Bobaljik, Jonathan. 2012. *Universals in Comparative Morphology: Suppletion, superlatives, and the structure of words*. MIT Press, Cambridge, Mass.
- Caha, Pavel. 2009. *The Nanosyntax of Case*. Ph.D. thesis, CASTL, University of Tromsø.
- Caha, Pavel. to appear. Explaining the structure of case paradigms through the mechanisms of Nanosyntax. *Natural Language and Linguistic Theory*.
- Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford University Press, New York.
- Cojocaru. nd. Romanian grammar. Available at www.seelrc.org/projects/grammars.ptml.
- Corbett, Greville G. 2000. *Number*. Cambridge University Press, Cambridge.
- Corbett, Greville G. and Richard J. Hayward. 1987. Gender and number in Bayso. *Lingua* 73: 1–28.
- D’Alessandro, Roberta and Ian Roberts. 2010. Past participle agreement in Abruzzese: Split auxiliary selection and the null-subject parameter. *Natural Language and Linguistic Theory* 28: 41–72.
- Feist, Timothy. 2010. *A grammar of Skolt Saami*. Ph.D. thesis, University of Manchester.
- Guillemin, Diana. 2011. *The syntax and semantics of a determiner system: a case study of Mauritian creole*. John Benjamins, Amsterdam.
- Ionin, Tania and Ora Matushansky. 2006. The composition of complex cardinals. *Journal of Semantics* 23: 315–360.
- Johnston, Jason Clift. 1996. *Systematic Homonymy and the Structure of Morphological Categories. Some Lessons from Paradigm Geometry*. Ph.D. thesis, University of Sydney.
- Lahne, Antje. 2007. On deriving polarity effects. In *1-2-many. One-to-many relations in grammar*, edited by Andreas Opitz and Jochen Trommer, no. 85 in *Linguistische Arbeits Berichte*, pp. 1–22. University of Leipzig, Leipzig.
- Manzini, Rita M. and Leonardo M. Savoia. 2011. Reducing ‘case’ to denotational primitives. *Nominal inflections in Albanian. Linguistic Variation* 11.1: 76–120.
- McCreight, Katherine and Catherine V. Chvany. 1991. Geometric representation of paradigms in a modular theory of grammar. In *Paradigms: The Economy of Inflection*, edited by Frans Plank, pp. 91 – 112. Mouton de Gruyter, Berlin, New York.
- Nevins, Andrew. 2011. Multiple agree with clitics: person complementarity vs. omnivorous number. *Natural language and linguistic theory* 29: 939–971.
- Nickel, Klaus Peter. 1990. *Samisk grammatikk*. Universitetsforlaget, Oslo.
- Pantcheva, Marina. 2010. The syntactic structure of locations, goals, and sources. *Linguistics* 48: 1043–1081.
- Pantcheva, Marina. 2011. *Decomposing Path. The nanosyntax of directional expressions*. Ph.D. thesis, CASTL, Tromsø.
- Plank, Frans. 1991. Rasmus Rask’s dilemma. In *Paradigms: The Economy of Inflection*, edited by Frans Plank, pp. 161–196. Mouton de Gruyter, Berlin.

- Plank, Frans. 1995. (Re-)introducing suffixaufnahme. In *Double Case: Agreement by Suffixaufnahme*, edited by Frans Plank, pp. 3–112. OUP, Oxford.
- Press, Ian. 2004. *Standard Lithuanian*. LINCOM, Muenchen.
- Richards, Norvin. 2007. Lardil “case stacking” and the structural/inherent case distinction. Ms., MIT, downloadable at: <http://ling.auf.net/lingBuzz/000405>.
- Starke, Michal. 2004. On the inexistence of specifiers and the nature of heads. In *Structures and Beyond: The Cartography of Syntactic Structures, vol. 3*, edited by Adriana Belletti, pp. 252–268. Oxford University Press, New York.
- Starke, Michal. 2009. Nanosyntax. A short primer to a new approach to language. In *Nordlyd 36: Special issue on Nanosyntax*, edited by Peter Svenonius, Gillian Ramchand, Michal Starke, and Tarald Taraldsen, pp. 1–6. University of Tromsø, Tromsø. Available at www.ub.uit.no/munin/nordlyd/.
- Starke, Michal. 2011. Towards elegant parameters: Language variation reduces to the size of lexically stored trees. Transcript from a talk at Barcelona Workshop on Linguistic Variation in the Minimalist Framework. Available at <http://ling.auf.net/lingBuzz/001183>.
- Stolz, Thomas, Cornelia Stroh, and Aina Urdze. 2006. *On Comitatives and Related Categories: A Typological Study with Special Focus on the Languages of Europe*. Mouton de Gruyter, Berlin New York.
- Stolz, Thomas, Cornelia Stroh, and Aina Urdze. 2009. Varieties of comitative. In *The Oxford Handbook of Case*, edited by Andrej Malchukov and Andrew Spencer, pp. 601–608. OUP, Oxford.
- Stump, Gregory. 2001. *Inflectional Morphology: A Theory of Paradigm Structure*. Cambridge University Press, Cambridge.
- Taraldsen, Tarald. 2010. The nanosyntax of Nguni noun class prefixes and concords. *Lingua* 120 6: 1522 – 1548.
- Trosterud, Trond. 2004. *Homonymy in the Uralic Two-Argument Agreement Paradigms*. Ph.D. thesis, University of Tromsø.
- Vangsnes, Øystein A. in press. Syncretism and functional expansion in Germanic wh-expressions. *Language Sciences* 34.
- Wiese, Bernd. 2003. Zur lateinischen Nominalflexion: Die Form-Funktions-Beziehung. Ms., IDS Mannheim.
- Wunderlich, Dieter. 2004. Is there any need for the concept of directional syncretism? In *Explorations in Nominal Inflection*, edited by Gereon Müller, Lutz Gunkel, and Gisela Zifonun, pp. 373–395. Mouton de Gruyter, Berlin.
- Wunderlich, Dieter. 2012. Polarity and constraints on paradigmatic distinctness. In *The morphology and phonology of exponence*, edited by Jochen Trommer, pp. 160–194. OUP, Oxford.