The feature structure of pronouns: a probe into multidimensional paradigms

Guido Vanden Wyngaerd
KU Leuven, CRISSP

Abstract

This paper examines multidimensional paradigms, i.e. paradigms involving more than one feature dimension. I will examine the concrete case of pronominal paradigms, which involve (at least) the dimensions of person and number. The problem that arises in such paradigms is that syncretisms may be observed in each dimension, i.e. they may occur both vertically (cross-person) and horizontally (cross-number). While classical nanosyntax embodies a theory of syncretism that can account for one dimension, it requires an extension to account for syncretisms in the other dimension(s). I discuss two such extensions, one making use of pointers (Caha & Pantcheva 2012), and another in terms of a revision of the Superset Principle (based on Caha 2014). I will show that both approaches make subtly different empirical predictions. Most notably, the approach in terms of pointers is able to derive certain ABA-syncretisms. This consideration favours the approach in terms of the revised Superset Principle. I conclude with a discussion of consequences of this revision for parts of the architecture of the nanosyntactic framework.

1 Introduction

Personal pronoun paradigms are multidimensional, as they involve (at least) the feature dimensions of person and number. Syncretisms in multidimensional paradigms may be horizontal and/or vertical:

(1)

	SG	PL
1	A	B
2	A	C
3	D	E

	SG	PL
1	A	A
2	B	C
3	D	E

As I demonstrate below, deriving the horizontal or cross-number syncretism is unproblematic given the Superset Principle, but deriving the cross-person or vertical syncretism requires an extension of the theory.

The structure of this paper is as follows: in sections 2 and 3, I lay out some assumptions I shall make concerning the internal structure of the person and number feature complex, respectively. Section 4 examines some attested horizontal and vertical syncretisms and shows how the horizontal ones can be derived and why the vertical ones are problematic. Section 5 examines the solution to this problem in terms of pointers, showing how they allow the derivation of ABA-patterns. Section 6 discusses another solution in terms of a revision of the Superset Principle. The different predictions these two approaches make are examined in detail in section 7. Section 8 addresses the issue of possible but unattested syncretisms. Finally, section **??** explores some further consequences of the revised Superset Principle.

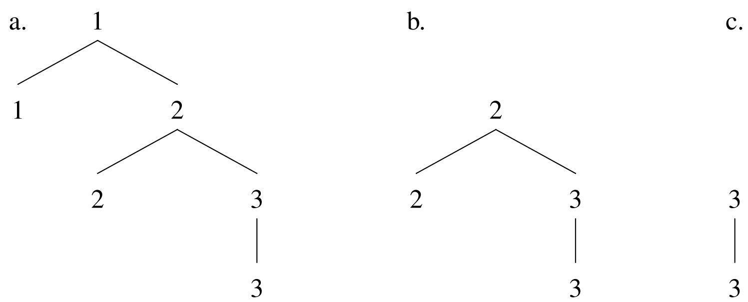
The empirical evidence on which this paper is based is mostly taken from the typological study by Cysouw (2003). I have nevertheless applied a number of restrictions to this material. Cysouw discusses all kinds of person marking, i.e. both independent pronouns as well as inflectional person markers. I focus exclusively on person marking in independent pronouns. This restriction is made possible by the fact that Cysouw is generally quite explicit on this issue in his description of the data.

An empirical domain that I shall leave aside in the present paper is the inclusive-exclusive distinction. Inclusive pronouns probably involve a composition of the features of first and second person pronouns, whereas exclusive ones only involve the features that enter into the makeup of the first person. However, since this is not the central topic of this paper, I leave a fully worked out nanosyntactic analysis as a topic for future research.

2 The person feature complex

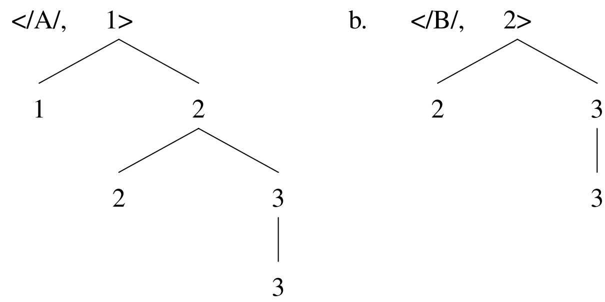
I adopt the proposal put forth by Starke (2013). This proposal assumes that there are three privative features are [speaker], [participant], and [person], which are moreover cumulative, i.e. they stand in a containment relationship. The first person personal pronoun has all three of these features, since the first person is the speaker, a participant and a person. The second person pronoun only has participant and person, and the third only person. For expository purposes, I refer to these features by numbers (1 = [speaker], 2 = [participant], and 3 = [person]). The feature trees for the first, second, and third person personal pronouns are given in (2a), (2b), and (2c), respectively:

(2)



Given these syntactic feature trees, we expect to find syncretisms between 1 and 2 (AAB), 2 and 3 (ABB), and 1, 2, and 3 (AAA), but no syncretism of 1 and 3 across 2 (*ABA). Consider a hypothetical example of an ABB-pattern, i.e. a 2/3 syncretism. This pattern can be shown to result from lexical items as given in (3):

(3) a. b.



In a 1 syntactic tree, (3a) will be the only candidate for insertion in virtue of the Superset Principle, since the lexical tree (3b) does not contain the syntactic tree as a subtree. In the second person, both lexical items are candidates, but (3b) will win the competition over (3a) in virtue of the Elsewhere Principle, as it applies in a proper subset of the environments in which (3a) applies. The same is true in the third person: both items are candidates for insertion but (3b) will win as it is a closer match.

3 Number

Before going on to a discussion of the actual syncretism data, we need to say something about the syntactic representation of number. Some languages form the plural of pronouns with the same morpheme that is used with nouns (or certain noun classes). This is illustrated for Mandarin Chinese in (4) (Corbett 2000:76):

(4)

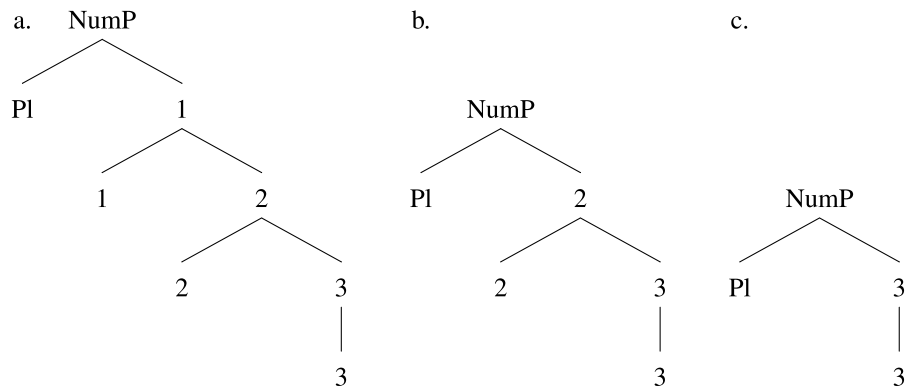
	SG	PL
1	wǒ	wǒ-men
2	nǐ	nǐ-men
3	tā	tā-men

(5) xuésheng xuésheng-men
student student-PL

Other languages displaying this pattern include Sierra Popoluca (Elson 1960), Trumai, Korean, Canela-Kraho, and Miskitu (Cysouw 2003, Corbett 2000; see Daniel 2013 for additional discussion).

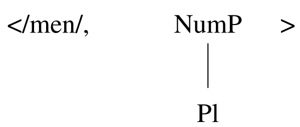
Exploiting the analogy illustrated in (4) and (5), I conclude that plural number sits on top of the person feature complex, as shown in (6):

(6)



In Mandarin Chinese, a lexical item as in (7) can be assumed, which spells out the plural morpheme:

(7)



Spell-out driven movement ensures that the complement ofPL in (6) moves into the Spec of NumP, after which *-men* spells out NumP. We shall have reason to modify the trees in (6) later, but for now they represent a good initial assumption on the position of the number dimension in the feature hierarchy.

4 Attested syncretisms

4.1 Types of patterns

Having outlined some of the background theoretical assumptions of this paper, we are now in a position to take a closer look at the types of syncretisms that one can observe in the personal pronouns. Basically, three general types of patterns can be distinguished, numbered I-III in (8).

(8)

	I		II		III	
	SG	PL	SG	PL	SG	PL
1	C	A	A	A	A	A
2	D	B	B	C	B	A
3	E	B	D	E	C	D

In the first pattern, one finds vertical, or cross-person, syncretisms. The second type of pattern instantiates horizontal, or cross-number, syncretism. The third type of pattern is a nonlinear one (i.e. cross-person and cross-number). All of these patterns are attested, and I shall now discuss them in turn, starting out with the horizontal syncretisms, as they are the simplest ones to derive.

4.2 Horizontal syncretisms

The horizontal syncretism can be restricted to a single person, two persons, or arise in all three persons. The following table lists the attested patterns:[1]

(9)

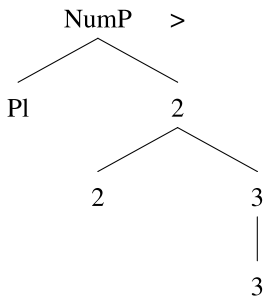
3	Sinhalese, Sentani, Asmat, SALISH
2 (rare)	English, Xokleng
1 (rare)	Marind
2 and 3	Berik, Kuman
1 and 3 (rare)	Tairora
all persons	Salt-Yui (3: demonstratives)

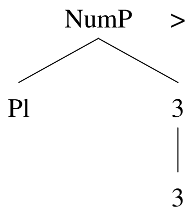
By way of example, consider the case of Berik (New Guinea):

(10)

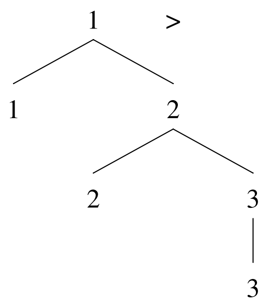
	SG	PL
1	ai	ne
2	aame	aame
3	je	je

The relevant lexical items are given in (11).

(11) a. </aame/,  >

b. </je/,  >

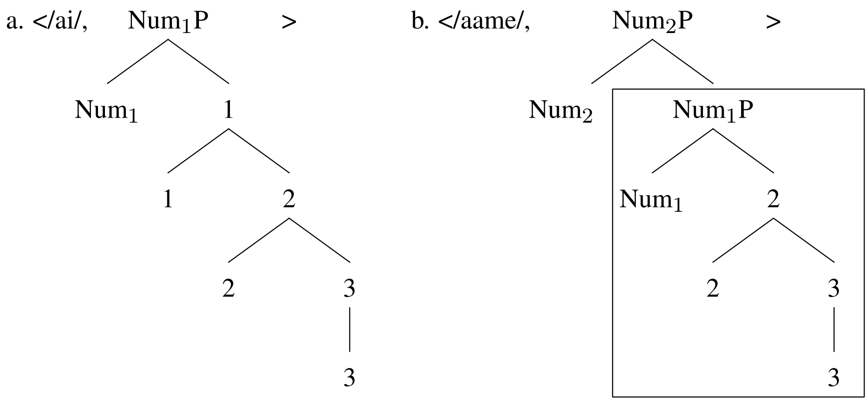
The item *aame* (11a) can spell out 2SG and 2L, by the Superset Principle: the tree of the singular pronoun is a subtree of the plural pronoun tree. For the same reason, (11b) can spell out 3, both singular and plural. A problem arises, however, when we consider the lexical item for the 1SG pronoun *ai*:

(12) </ai/,  >

[1]Small capitals indicate language families rather than individual languages. Quite a number of languages feature no third person pronouns, but replace them by demonstratives instead. I shall leave it as an open question whether this state of affairs is to be considered an instance of a lexical gap, filled by the demonstratives, or whether we are dealing with a syncretism between third person personal pronouns and demonstratives. The systematic nature of the filling of the gap by demonstratives suggests that the latter interpretation is correct.

The problem is that in 2SG, there is a tie between (11a) and (12): (11a) *aame* and (12) *ai* each contain exactly one feature more than the syntactic node of a 2SG pronoun. This problem can be solved by assuming that the number projection is internally complex, i.e. singular number also involves the presence of a number feature (Num_1), plural number involves two features (Num_2 and Num_1).[2]

(13)



Now (13b) *aame* can still spell out 2, both singular and plural: in the plural it is a perfect match, and in the singular the lexical tree contains the syntactic tree as a subtree (highlighted in (13b)). However, (13a) *ai* can no longer spell out the 2SG, since it does not contain the syntactic tree as a subtree. This solution crucially requires that singular pronouns contain a Num_1 feature: the presence of Num_1 in (13a) prevents the tree from being a match for a 2 syntactic tree. Informally, one can describe this as an impossibility of a lexical tree to 'shrink' in the middle to match a syntactic tree. While trees can 'shrink' at the top to match a syntactic tree, they cannot shrink in the middle. The other attested patterns of horizontal syncretism work in the same way, and can be derived in the same way.

Summarizing, the horizontal syncretisms support the claim that singular number is not the absence of number, but the presence of a singular number feature. The existence of horizontal syncretisms further rests on the possibility to build trees with an incomplete person f_{seq}, i.e. with person features missing at the top of the person sequence, as well as the possibilty of 'shrinking' the number projection at the top of the tree, as allowed by the Superset Principle.[3]

4.3 Vertical syncretisms

Cross-person syncretisms in the singular pronouns are extremely rare: Cysouw (2003) finds only two languages (out of the 90 for which independent pronoun paradigms are

[2] I leave open the question of the precise nature of these number heads. A. Rocquet (p.c.) suggests that the first number head (Num_1), which is always present in the pronouns, could be a layer of 'definiteness' or a 'count' layer, since personal pronouns usually refer to definite/count entities. The number projection is very likely to be more complex than suggested here, as it does not make provisions for the well-known property of dual number. Since the analysis of the number projection is not the topic of this paper, I shall leave this as an issue for further research.

[3] The metaphor of tree shrinking used here is unrelated to the concept of impoverishment in DM. Rather, it refers to the fact that a lexical tree may map onto a syntactic tree that is smaller, i.e. the lexical tree may be overspecified in comparison with the syntactic tree; as such, it is the nanosyntactic counterpart of the Subset Principle of DM.

listed) showing ABB or AAB, Qawesqar and Winnebago.[4] Vertical or cross-person syncretisms do occur in the plural, however. Some languages that display them are listed below:

(14) a. AAB: many Athabascan languages (e.g. Slave, Chiricahua Apache, Navaho, Kato, Hupa), Awa, Southern Haitian Creole

b. ABB: Nez Perce, Warekena, Wolof (object pronouns), Mauritian Creole[5]

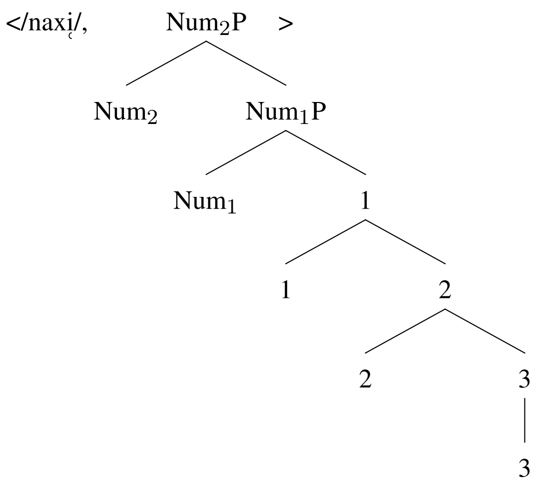
The account of the AAB and ABB syncretisms is not straightforward. Consider the AAB-pattern in Slave (Cysouw 2003:124):

(15)

	SG	PL
1	sį	naxį
2	nį	naxį
3	ʔedį	ʔegedį

The lexical tree for the 1PL pronoun is given in (16):

(16)



This can spell out the syntactic tree of a 1PL pronoun, but not of a 2PL one, since the latter lacks the 1 node and is therefore not a subtree of (16). To derive the AAB-pattern, the tree would have to be able to shrink in the middle (from 1 to 2). For the same reason, the ABB-pattern cannot be derived either, as the lexical item for 2 cannot shrink to 3.

A solution to the problem of the existence of both horizontal and vertical syncretisms in multidimensional paradigms has been proposed by Caha & Pantcheva (2012), which relies on the mechanism of pointers. A different solution relies on a reformulation of the Superset Principle, which can be applied to the problem of multidimensional paradigms as well. I shall now discuss these two solutions in turn.

[4]Concerning Qawasqar, however, M. Cysouw (p.c.) informs me that work by Oscar Aguilera reveals that the data in Clairis (1985) are probably not entirely reliable, and that Qawasqar has different forms for 1SG and 2SG, and either a null pronoun or a demonstrative in 3SG. A language not mentioned by Cysouw is Sanapaná (Gomes 2013), which has an ABB pattern both in the singular and the plural.

[5]According to Baker (1972) and Stein (1984), but not Adone (1994), who gives an ABC-pattern in the plural.

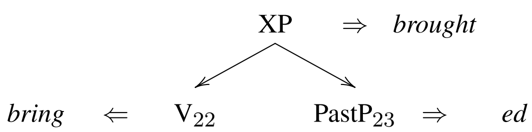
5 Pointers

5.1 Suppletion

A pointer is a node in the tree of a lexical item that points to another, existing, lexical item (Starke 2011). Starke proposes pointers as a mechanism to account for cases of suppletion.[6] For example, the lexical item of the suppletive past tense *brought* contains a pointer (represented by the numerical indices in (17a)) to the lexical items for *bring* on the one hand, and the past tense morpheme on the other.

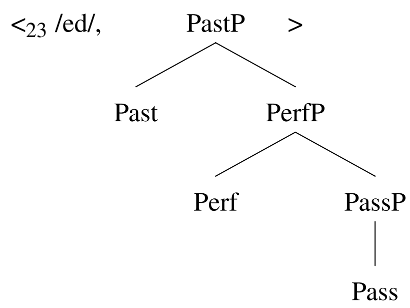
(17) a. <$_{24}$ /brought/, [$_{XP}$ 22 23]>
b. <$_{22}$ /bring/, V>
c. <$_{23}$ /ed/, PastP>

Informally, the pointer may be represented in a tree using an arrow instead of aPLain branch:

(18)


Each of the lexical items pointed to is subject to independent cyclic spellout. This creates *bring+ed*, which is overwritten at the top node by *brought*. Given the syncretism between Past-Perfect-Passive illustrated in (19), we must conclude that *-ed* has more internal structure, so that instead of (17c), we have (20):

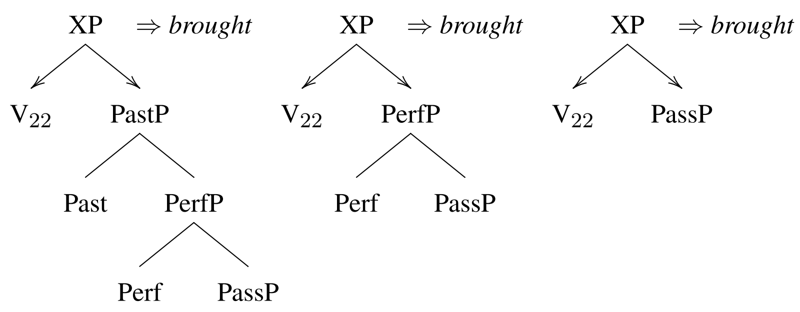
(19) a. They elected George.
b. They have elected George.
c. George was elected.

(20)


The suppletive form *brought* shows the same Past-Perfect-Passive syncretism. This means that in the item with the pointer (19), the item pointed to (20) must be allowed to shrink to any subtree.

[6] Starke (2014) presents a slightly different take on suppletion, but not one that has different effects, as far as I can tell.

(21)

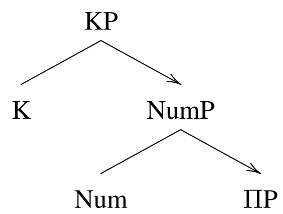


This is a natural consequence of the mechanism of cyclic spellout: the item pointed to is an independent lexical item and can match a syntactic tree that the lexical tree contains as a subtree. Informally put, an item with a pointer can shrink not just at the top, but also in the middle of the tree, at the top of the item pointed to. As a result, the lexical item *brought* can spell out three different syntactic trees, as illustrated in (21).

5.2 Multidimensional paradigms

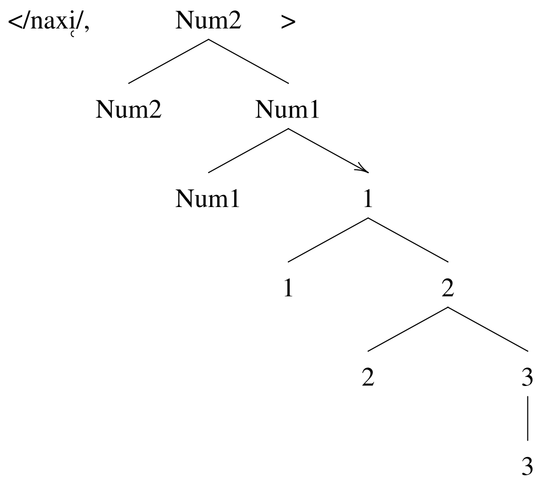
Let us return now to the vertical syncretisms in the pronoun paradigms in the plural. As we saw above, pronouns spell out multiple feature dimensions: Case, number, person, and gender. Let us assume that the lexical items for pronouns can also contain pointers at the juncture of the dimensions (in the tree below, KP represents the Case sequence, NumP the number sequence, and ΠP the person sequence):

(22)



This assumption allows the derivation of the problematic vertical syncretisms, since the tree can now shrink in the middle (from 1 to 2 to 3). Recall the lexical tree for the Slave pronoun *naxį*, syncretic for 1PL and 2PL (see (16) above). We now add a pointer to this tree:

(23)



This representation allows us to derive the AAB-pattern. The lexical item in (23) can spell out a 1PL pronoun, but also a 2PL one, because of the presence of the pointer. The lexical item for the 3PL pronoun *ʔegedi̢* does not contain the 1 and 2 projection. As a result, it will win the competiton from (23) in the 3PL because of the Elsewhere Principle.

In a similar fashion, vertical ABB syncretisms can be derived. Assume a lexical item like (22) but without a pointer. This is the A-pronoun, which can only spell out 1PL, since it does not contain a pointer. In addition, assume a B-pronoun like (22) (with a pointer) but without the 1 node. This B-pronoun does not compete with the A-pronoun in the first person, since it lacks the 1 node. The B-pronoun contains a pointer and can spell out both 2PL and 3PL, winning against the A-pronoun because of the Elsewhere Principle.

5.3 Pointers introduce ABA

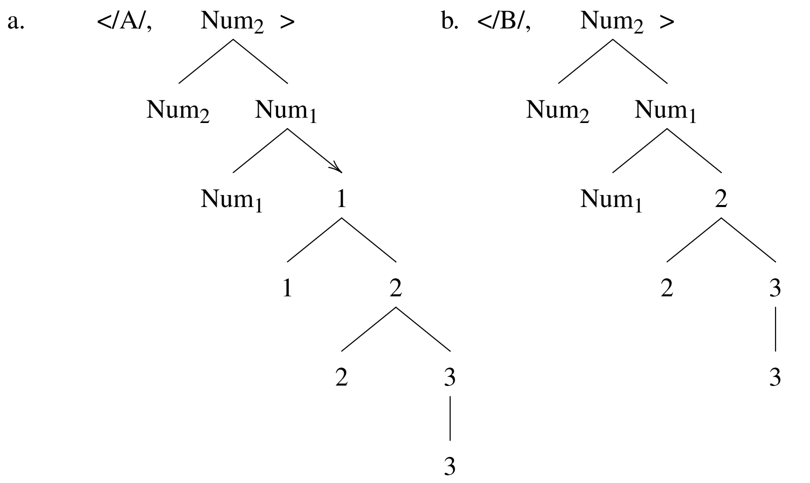
Allowing pointers also allows a certain type of ABA-pattern in the plural, in agreement with an abstract prediction made by Taraldsen (2012). Consider (24):

(24)

	SG	PL
1	C	A
2	D	B
3	E	A

Let us see how the derivation of this pattern is allowed from the system outlined above. For A, assume a lexical entry with a pointer, as in (25a), which is flexible at the number-person joint: due to the shrinkability of items with pointers, the lexical item A can spell out all the plural pronouns. The lexical item for B, given in (25b), does not contain a pointer, and is therefore rigid (i.e. not shrinkable in the middle).

(25)



If the syntactic tree is 3PL, (25a) A is the only candidate, since (25b) cannot shrink in the middle to spell out 3PL. If the syntactic tree is 2PL, B wins the competition from A because of the Elsewhere Principle. This is because the lexical item (24b), without the pointer, applies in a proper subset of the environments of the lexical item (25a), with the pointer: (25a) can shrink both at the top and in the middle and therefore applies to 9 structures, whereas (25b) can only shrink at the top and applies to only 4 structures. If the syntactic tree is 1PL, B is not a competitor since it lacks a 1 node; A can (and does) spell out the tree. These findings agree with an abstract prediction made by Taraldsen (2012), who argues that ABA-patterns may arise in multidimensional paradigms, given Caha & Pantcheva's analysis in terms of pointers.

In sum, we have seen that vertical syncretisms in the pronouns can be derived by means of assuming the presence of pointers in lexical representations. But assuming pointers opens the door to ABA-patterns. Given the central role that syncretisms play in establishing feature hierarchies, this is a potentially undesirable conclusion.[7] I now therefore turn to another solution, which involves a revision of the Superset Principle, and which does not allow the derivation of ABA-patterns.

[7] It is an open question whether ABA is actually attested in the independent pronouns. I am aware of two claims in the literature to that effect. Baerman et al. (2005) quote the case of Dakar Wolof (Nussbaum et al. 1970) as an example of a language with a 1/3 syncretism in the independent pronouns. However, this case is moot. Stewart & Gage (1970:38) give the following paradigm for the Dakar Wolof independent pronouns, which does not show the 1/3 syncretism:

(i)

	SG	PL
1	man	ñum
2	yow	yeen
3	moon	ñoom

There is, to be fair, a 1/3 syncretism in the plural of the so-called dependent pronouns, but these behave more like clitics than like independent pronouns, to the point where one could wonder if they had not better be analysed as inflectional markers. The second case involves Bagirmi, a Nilo-Saharan language spoken in Chad (Gaden 1909), which Cysouw (2003) claims has an ABA-pattern in the plural pronouns. This does, however, appear to be an isolated case, from a relatively undocumented language.

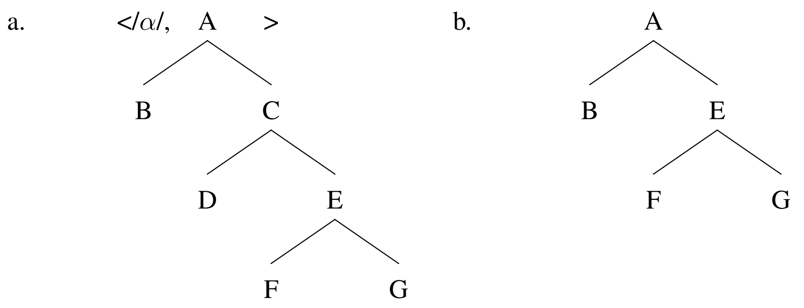
6 Reformulating the Superset Principle

In this section, I propose a reformulation of the Superset Principle, which will allow us to capture both horizontal and vertical syncretisms in multidimensional paradigms, but which at the same time does not make ABA derivable. It relies on a reformulation of the Superset Principle, which is inspired by (but not identical to) a proposal made by Caha (2014).

(26) *Revised Superset Principle (RSP)*
A a lexical entry L may spell out a syntactic node SN iff the features of L are equal to or a superset of the features dominated by SN.

The most striking difference between the RSP and the classical Superset Principle is the absence in the RSP of a reference to a subtree of the lexical tree as a condition on the spellout of syntactic trees. The RSP replaces this requirement with a weaker superset condition on terminals or features of the tree. This reformulation will allow the 'shrinking in the middle' that was needed to derive the vertical syncretisms in pronominal paradigms. Consider the following abstract example, where (27a) is a lexical item and (27b) a syntactic tree:

(27)



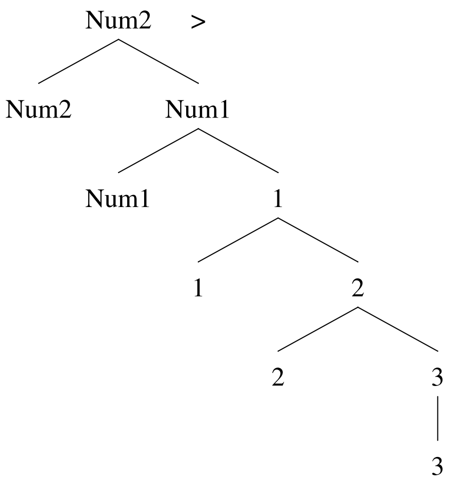
The RSP makes (27a) a possible spellout for the syntactic tree (27b), even though the syntactic tree misses the C and D nodes that occur in the middle of the lexical tree. All the RSP requires is that the features of the lexical item (which are the terminals {B, D, F, G}) be a superset of the features of the syntactic tree (which are {B, F, G}. This condition is obviously met in (27). Put more informally, the lexical tree can 'shrink' at any point to become identical to the syntactic tree.

The RSP derives both horizontal and vertical syncretisms in multidimensional paradigms without the need for pointers, and at the same time makes ABA-patterns underivable. To see how this works, reconsider the AAB-pattern in Slave, with the lexical tree for the 1PL pronoun given in (29):

(28)

	SG	PL
1	sį	naxį
2	nį	naxį
3	ʔedį	ʔegedį

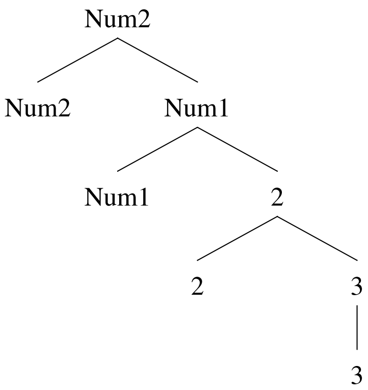
(29) </naxi̜/,



>

The lexical item (29) will be able spell out the syntactic tree for 1PL, as both trees are fully identical. What we need to show is that (29) can also spell out the syntactic tree for 2PL, given the RSP. The relevant syntactic tree is given in (30):

(30)



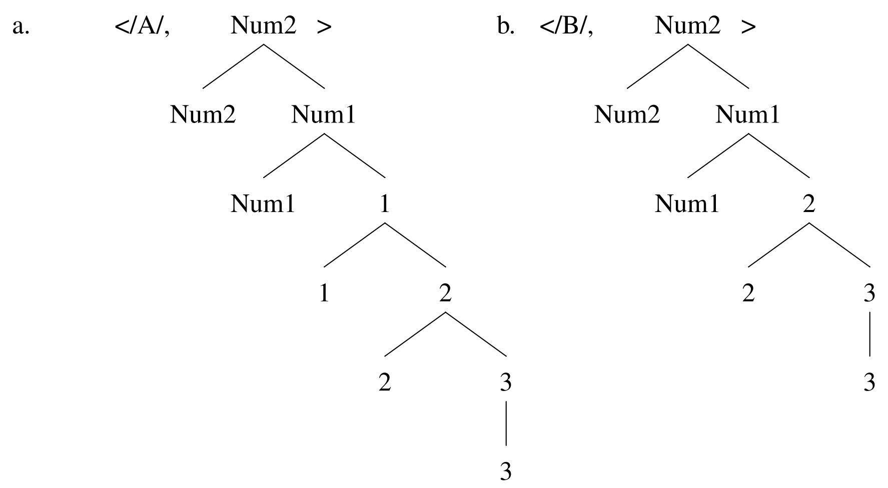
The lexical item (29) can spell out the syntactic tree (30), given that the its features (Num_2, Num_1, 1, 2, 3) are a superset of those of the syntactic tree (Num_2, Num_1, 2, 3). Note that the lexical tree does not contain the syntactic tree as a subtree: in this example, the node 1 of the lexical tree is not found in the syntactic tree. While this situation is ruled out by the classical Superset Principle, it is allowed by the RSP.

The RSP does not allow the derivation of ABA-patterns. Recall the ABA-pattern discussed above, and the corresponding lexical items in (32) (but now without a pointer).

(31)

	SG	PL
1	C	A
2	D	B
3	E	A

(32)



Given the RSP, both A and B can shrink in the middle, i.e. both can now spell out 2 and 3. As a result, both lexical items will compete in 2PL and 3PL. The Elsewhere Principle now ensures that B will win the competition both in 2PL and 3PL, as B contains less superfluous features. As a result, *ABA holds in full generality, i.e. ABA-patterns are underivable in principle.

In sum, both the approach in terms of pointers and the RSP allow the derivation of multidimensional paradigms. However, the pointers approach opens the door to the derivation of ABA-patterns. The RSP is more restrictive and does not allow the derivation of ABA-patterns. In what follows, I investigate a series of nonlinear syncretisms, showing where both approaches make different predictions.

7 Nonlinear syncretisms

7.1 Shapes and sizes

In this section, I discuss syncretisms which are not exclusively horizontal, and not exclusively vertical either. These come in a variety of shapes and sizes, which are summed up in (33):

(33) a. L-shaped, contiguous
b. double L, without ABA
c. diagonal (non-contiguous)
d. L-shaped, with ABA
e. double L, with ABA

The interest of the discussion is that the two approaches for dealing with multidimensional paradigms outlined above make different predictions as to the possible existence of these patterns. In what follows, attested patterns are illustrated by an example, and unattested patterns by an abstract representation of the syncretism type in terms of capital letters.

7.2 L-shaped, contiguous

There basically four different logically possible L-shapes, only one of which is attested. This is the type of L-that is found in Usarufa (Cysouw 2003):

(34)

	SG	PL
1	ke	ke
2	e	ke
3	we	ye

This syncretism is derivable under the pointers approach: the 1 lexical item *ke* contains a pointer, which means that it can spell out all persons and numbers. However, *ke* loses the competition to more specific lexical items without pointers (*e, we, ye*).

The pattern is also derivable with the RSP: here too, *ke* can spell out all persons and numbers. In 2SG, *ke* loses the competition to the more specific lexical item *e*, which lacks the 1 node. In 2PL, *ke* is the only candidate, because *e* lacks the plural number node, and *we* and *ye* lack the 2 node. In 3, *we* and *ye* win from all the other items because they lack the 2 and 1 nodes, and therefore have less superfluous features.

The second type of L is one that is a mirror symmetry of the first one:

(35)

	SG	PL
1	A	A
2	A	B
3	C	D

This pattern is also derivable under the pointers approach: the A-item can spell out all persons and numbers because it can spell out 1PL and because it contains a pointer. It will lose against the more specific B, C, and D items where these occur because these do not contain a pointer. In the RSP approach, however, this pattern is underivable. The A-item is maximal and can spell out all persons and numbers. The B-item can spell out anything that is nonfirst person. As a result, the Elsewhere Principle will yield a win of the B-item over the A-item in 2SG and 2PL. This makes the pattern in (35) effectively underivable.

Both of these patterns have a variant where the L-shape is one row lower in the paradigm. Our conclusions on (34) and (35) extend directly to these variants, so that we shall not discuss them in detail here.

Two other types of L-shaped patterns involve a 180° rotation of the two previous patterns:

(36)

	SG	PL	SG	PL
1	A	B	B	A
2	A	A	A	A
3	C	D	C	D

These are both derivable under the pointers approach, for essentially the same reason as above: A can spell out all persons and numbers and contains a pointer, but loses from B, C and D where these occur because these do not contain pointers. Neither pattern is derivable under the RSP approach, however. The left hand one has an A-item that can spell out all persons and numbers since it can spell out 1 and plural number. But the B-item can do so, too, as it spells out 1PL. As a result, there will be a tie between A and B in all cells of the paradigm that are nonthird person. The paradigm on the right of (36) is underivable because A cannot win from B in 2SG, as B is more specific in never occurring in the plural. The same conclusions hold for the variants of these patterns where the L is one row lower in the paradigm.

In sum, of the four logically possible L-shapes (and their variants), only one is actually attested. This is also the only one derivable under the RSP, whereas all four types are derivable under the pointers approach. The RSP approach therefore provides a much better fit with the data here than the pointers approach.

7.3 Double L, without ABA

This abstract pattern is illustrated in (37) and (38):

(37)

	SG	PL
1	A	A
2	B	A
3	B	B

(38)

	SG	PL
1	A	A
2	A	B
3	B	B

Neither of these patterns is derivable under the pointers approach. The reason is that there are two competing items, which both contain pointers. The A-item is maximal and flexible, and can express all persons and numbers. The B-item is more specific, since it does not contain a 1 projection, but it is also flexible, and can apply to both 2 and 3 and to both siingular and plural. This will ensure a win of A in the first person, but also a win of B in all the other persons. In particular, in the second person the B-item will win both in the singular and in the plural, because the A-item applies to more cases than the B-item.

The same reasoning applies to the approach in terms of the RSP. Since B can express both singular and plural number, it will win against A both in the singular and in the plural.

7.4 Diagonal

Diagonal syncretisms contradict spatial accounts of syncretism (e.g. McCreight & Chvany 1991). As we shall see, there are quite a number of logically possible types of diagonal syncretisms, but only one of these types is attested. This is the pattern illustrated by the following example from Suki:

(39)

	SG	PL
1	ne	e
2	e	de
3	u	i

This pattern is derivable under the pointers approach. The lexical tree of the *e*-pronoun is maximal and flexible, i.e. shrinkable at the joint (from 1 to 2). As a result, *e* can express all the persons and numbers, but it loses the competition against the rigid items for the other persons and numbers, since they contain less superfluous material.

Diagonal syncretisms of this type are underivable with the RSP. Here too, the lexical item *e* can spell out all persons and all numbers. In 2SG, there are two more specific items, however: *ne* (because it cannot spell out plural) and *de* (because it cannot spell out first person). The competition between *ne* and *de* will be won by *ne*, because *ne* can spell out 1/2/3SG; *de* can spell out 2/3SG and 2/3PL, so *ne* is more specific than *de*.

Now look at another type of diagonal, which is minimally different from (39), but which is unattested:

(40)

	SG	PL
1	A	B
2	C	A
3	D	E

This type of diagonal is also derivable with pointers (for the same reason as (39)). However, it is underivable with the RSP. Both the A and the B-item can spell out all numbers and all persons: B because it spells out the maximal tree of 1PL, and can consequently spell ou any smaller tree. Likewise, A can spell out any person or number, as it can spell out 1 (the maximal person) as well as plural (the maximal number). As a result, the lexical trees of A and B would be identical. This would result in a tie between A and B in 1SG and 1PL), and in 2PL. I take this to be an inadmissible situation in the domain of pronouns.

Two other types of diagonal patterns are also unattested (for the sake of brevity, I show the distinct syncretism patterns in a single table):

(41)

	SG	PL	SG	PL
1	B	C	B	C
2	D	A	A	D
3	A	E	E	A

As before, A contains a pointer and is therefore flexible in the middle, but it will lose to the nonflexible items D and E exactly where it has to. The RSP approach does not allow the derivation of these patterns. The two diagonals in (41) have basically the same shape as the first two in this section, and they are ruled out for the same reasons: in the pattern on the left, E will win in 3SG, in that on the right, there will be a tie between A and D in 2 and 3PL.

A combination of both diagonals yields the double diagonal patterns in (42), both unattested but derivable with pointers, though not with the RSP:

(42)

	SG	PL	SG	PL
1	A	B	B	A
2	C	A	A	C
3	A	D	D	A

Summarizing so far, there are six types of diagonal syncretisms allowed by the pointers approach, only one of which is actually attested. All types of diagonals are ruled out by the RSP.

We now proceed with a discussion of some further types of diagonals, which combine the diagonal with an L-shaped syncretism. These syncretisms are underivable under either approach. An overview of the possible patterns is given in (43):

(43)

	SG	PL	SG	PL	SG	PL	SG	PL
1	C	A	A	C	B	B	B	B
2	A	B	B	A	B	A	A	B
3	B	B	B	B	A	C	C	A

These are underivable with pointers because both A and B contain a pointer. The problem arises in the rows where A competes with B in the same row: either B wins the competition both in the singular and the plural (as in the two patterns on the left), or A wins both in the singular and the plural (as in the two patterns on the right). The patterns are also underivable with the RSP, as are all patterns involving diagonals.

There are some further logically possible diagonal syncretisms, which combine a diagonal syncretism with an ABA pattern, as illustrated in (44):

(44)

	SG	PL
1	C	A
2	D	B
3	B	A

This is not derivable under the pointers approach, again because both A and B must be assumed to contain pointers, both the vertical and the diagonal syncretism requiring pointers, as explained above. As a result, B will win from A in the 3PL. Neither are syncretisms with ABA derivable under the RSP, for the reasons explained above. There are more logically possible patterns which combine a diagonal (or a double diagonal as in (42)) with an ABA (or a double ABA), but since they are all underivable we shall not discuss them any further here.

7.5 L-shaped, with ABA

An L-shaped syncretism with ABA would be one instantiating the abstract pattern in (45):

(45)

	SG	PL
1	A	A
2	C	B
3	D	A

The syncretism is derivable in the pointers approach: the A-item is maximal and flexible, and it loses out to the more specific C-B-D items where the latter occur. The pattern is underivable with the RSP, as are all cases involving ABA, for the reasons outlined earlier.

7.6 Double L, with ABA

One instance of this abstract pattern is illustrated in (46):

(46)

	SG	PL
1	A	A
2	B	B
3	B	A

This is not derivable under the pointers approach for the same reason as the other double L patterns discussed above: there are two competing items, both containing pointers, with the tree for B smaller (and therefore more specific) than the tree for A. This will ensure a win for A in 1, as required. The problematic cell is 3PL, where B wins against A since its tree is smaller than the tree of A. The pattern is also underivable under the approach in terms of the RSP, since it involves an ABA pattern in the plural.

7.7 Summary of findings

Table 1 presents an overview of the findings in this section. The top line represents the syncretism patterns by their example number; a superscripted number indicates the number of different syncretism patterns the example instantiates, if there is more than one. The second line indicates whether a pattern is derivable with pointers, and the third if the RSP allows it. The bottom line indicates whether a pattern is attested.

Table 1: Overview of derivable and attested syncretisms

	34	35	36^2	37	38	39	40	41^2	42^2	43^4	44	45	46
PTR	✓	✓	✓	✗	✗	✓	✓	✓	✓	✗	✗	✓	✗
RSP	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗	✗
ATT	✓	✗	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗

What the table shows is that both the approach in terms of pointers and the one in terms of the Revised Superset Principle face some empirical problems. The RSP approach is more restrictive in general, but perhaps too strongly so, in that it rules out the existence of a pattern which is actually attested. This is the diagonal pattern found in Suki (i.e. (39)). According to Cysouw (2003:121), this pattern is 'commonly found in the contemporary Aztecan languages', although the evidence he quotes involves subject prefixes rather than independent pronouns. On the other hand, the pointers approach may be said to be too liberal, in that it allows far more syncretisms than are actually attested. This is particularly true in the diagonal syncretisms, where only one type is attested, but many more allowed by the pointers approach. The situation can be represented in a subset-superset subset hierarchy, where neither the approach in terms of the RSP nor the one in terms of pointers provides an exact match with the attested data.

(47) logically possible $\supset$ derivable with pointers $\supset$ attested $\supset$ derivable with RSP

On conceptual grounds, the approach in terms of the RSP seems preferable, as it rules out all cases of ABA, and it therefore leaves the syncretism diagnostic, a cornerstone of the nanosyntactic method, fully intact. At the same time, it remains an empirical question whether in certain cases ABA-patterns are attested or not. The empirical domain investigated here is that of the independent pronouns, where they are indeed vanishingly rare, but in the domain of verb agreement the existence of 1-3 syncretisms is uncontroversial (see e.g. Ackema & Neeleman 2013). Note further that even under the pointers approach, ABA-patterns could not arise in the dimension that is structurally highest. In the nominal domain, this is arguably the Case dimension. Even if a lexical item could contain a pointer in the Case sequence, this would not lead to any ABA-patterns in the Case dimension. It would only create the possibility for spelling out sequences that the syntax cannot create, assuming there to be no gaps in the functional sequence for Case. The possibility of ABA is only created in dimensions that sit lower in the hierarchy than Case, such as number and person.

8 Possible but unattested syncretisms

In this section, I want to briefly address the matter of possible but unattested syncretisms. There is a rather striking such case in the data discussed so far, which is that of the vertical syncretisms in the singular. The fact that such syncretism is attested in the plural suggests that the theory should allow it, yet it is virtually unattested in the singular. The question is why this should be so. A possible answer comes from an observation by Baerman et al. (2005:44), who observe that case syncretisms are sensitive to an animacy hierarchy, given in (48):

(48) 1sg > 2sg > 1incl du > 1inclPL > 2du > 2l > 3 > this > that > indefinite > animate > *meat, vegetable* > other inanimate

What they note is that in nominative/accusative case systems, NOM-ACC syncretisms are more likely to occur towards the right-hand side of the hierarchy, i.e. in the less animate nominals. For instance, Indo-European languages regularly have NOM-ACC syncretism in nouns but typically lack it in pronouns. The systematic NOM-ACC syncretism found in neuter nouns in Indo-European (e.g. Greek and Latin) can also be seen as an instance of this correlation, in so far as these nouns typically have inanimate referents. In languages with split ergativity, on the other hand, the correlation is opposite, i.e. ERG-ABS syncretisms are more likely to occur with items higher on the hierarchy.

The hierarchy in (48) in fact conflates a number different subhierarchies:

(49)
a. Person: 1 > 2 > 3
b. Number:SG > du >PL
c. Animacy: animate > inanimate
d. Definiteness: definite > indefinite
e. Word class: pronoun > demonstrative > noun

Given the number subhierarchy (sg > non-singular), one expects this correlation to also be found with respect to number, i.e. non-singular number should favour NOM-ACC syncretism, and singular number ERG-ABS syncretism. However, this prediction is not borne out. Instead, both patterns of syncretism are more likely in the non-singular than in the singular. Similarly, gender syncretism occurs more easily in the non-singular than in the singular, as in e.g. Germanic, where gender distinctions are typically lost in the plural. Obviously, this tendency for syncretism to occur more easily in the non-singular than in the singular strikingly resembles the one we observe in the pronoun paradigms under discussion here. Another reflex of the correlation between animacy and syncretism can be found in the fact that horizontal (cross-number) syncretism is far more common in the third person than in the first and second persons (see (9) above). These observations do not in and of themselves constitute an explanation for derivable but unattested (or rarer) syncretisms, but they do point to more general patterns that can be observed in other domains as well.

9 Further consequences

This final section discusses both a problem and a consequence of the RSP. The problem concerns horizontal syncretisms in the nonfirst person. The abstract pattern is given in (50):

(50)

	SG	PL
1	A	B
2	C	C
3	D	E

Here, both the lexical items A and C are candidates for spelling out a 2SG syntactic tree, but the Elsewhere Principle does not designate either one as a winner. C can spell out 2SG because it can spell out 2PL, i.e. it contains exactly one more feature than the syntactic tree. The same is true for A, however: assuming either pointers or the RSP, A can spell out 1SG and therefore also 2SG; like C, the lexical item for A contains exactly one feature more than the syntactic tree. Since A and C contain an equal amount of superfluous material, the Elsewhere Principle cannot determine a winner, an unwelcome result. This is in fact a problem that we encountered earlier (in connection with the horizontal syncretism of Berik discussed in section 4.2 above), and that we solved by assuming that singular number involved a separate number feature, rather than the absence of number. This solution worked as long as we only considered horizontal syncretisms, i.e. trees that only could shrink at the top. But now that we have introduced mechanisms that allow trees to shrink in the middle as well, the problem reappears. In the pointers approach, there is an easy way to solve it: the only assumption needed is that A contains no pointer, and therefore is no candidate for spelling out 2SG. C only has to shrink at the top and will therefore spell out 2SG. In the RSP, however, an additional assumption needs to be made to ensure that C is the winner. This follows from the fact that the RSP gives up the subtree requirement that is present in the classical Superset Principle, and only looks at the features. Since C and A both have exactly one feature more than the syntactic tree, they are both candidates, and both contain exactly the same amount of superfluous material. A solution consists in claiming that the Elsewhere Principle must be supplemented with the following principle:

(51) Prefer spellouts where the mismatch between the lexical tree and the syntactic tree is restricted to the top of the tree.

An informal version of this principle could be formulated as 'Prefer High Junk', or 'Minimize Gaps'. This will ensure a win of C against A in 2SG: C only differs from the syntactic tree at the top, whereas A differs from it in the middle. In sum, in order to derive horizontal syncretisms, the RSP needs to be supplemented with the principle in (51).

The second point to be discussed in this section concerns a consequence of the RSP for the way the lexicon is organized. The most striking difference between the RSP and the classical the Superset Principle is the absence in the RSP of a reference to a subtree of the lexical tree as a condition on the spellout of syntactic trees. Since the RSP gives up the concept of a subtree, an important piece of motivation is lost for the assumption that lexical items contain trees. The alternative would be to assume that lexical items simply contain sets of features. To see what this would look like, reconsider the abstract example (27) discussed above, where the syntactic tree contained the terminals B, F, G, and the lexical entry an additional terminal D between B and F. An alternative lexical entry for (27a) would therefore contain a mere unordered set of terminals (or features), as shown in (52):

(52) </α/, {B, D, F, G}>

This will be insertable in a syntactic tree with the terminals {B, F, G} by the RSP. Obviously, some information is lost in a representation like (52) as compared with the richer kinds of representations with lexical trees. But what gets lost is largely redundant information. For one thing, we lose the nodes that are the projections of these heads in the tree; but this is redundant information, as their presence follows from the general principle that each feature is a syntactic head that projects. For another, we lose the hierarchical order of these heads, but this is also easily recoverable from the universal functional sequence. Since it is not clear at this point whether lexical representations as in (52) are more than mere notational variants of the lexical representations containing trees, I shall not dwell on the matter any further here.

10 Conclusion

I have shown that the analysis of syncretism in multidimensional paradigms requires an extension of classical nanosyntactic theory. I have discussed two such extensions, one in terms of pointers, and another in terms of a Revised Superset Principle. Both approaches make different empirical predictions, notably with respect to the possible existence of ABA-patterns, and of certain types of diagonal syncretisms. Empirically, neither approach provides a perfect fit with the available data. At the conceptual level, the RSP, though ruling out any kind of ABA in principle, requires an additional principle over and above the Elsewhere Principle to derive horizontal syncretisms in nonfirst persons.

Acknowledgements

I am grateful to the anonymous revieuwers of OUP for their feedback. For useful discussion and suggestions, I wish to thank Pavel Caha, Karen De Clercq, Michal Starke, and the audiences at SLE 47 in Poznan and at SinFonIJA9 in Brno.

References

Ackema, Peter & Ad Neeleman. 2013. Person features and syncretism. *Natural Language and Linguistic Theory* 31. 901–950.

Adone, Dany. 1994. *The acquisition of Mauritian Creole* Amsterdam. John Benjamins Publishing Company.

Baerman, Matthew, Dunstan Brown & Greville Corbett. 2005. *The syntax-morphology interface: A study of syncretism*. Cambridge: Cambridge University Press.

Baker, Philipp. 1972. *Kreol*. London: C. Hirst and Company.

Caha, Pavel. 2014. Funky genitives. Class lectures, GLOW Spring School, Brussels.

Caha, Pavel & Marina Pantcheva. 2012. Contiguity beyond linearity. Talk at Decennium: The first 10 years of CASTL.

Clairis, Christos. 1985. *El qawesqar: lingüística fueguina, teoría y descripción*, vol. 12 Estudios Filológicos, Anejo. Valdivia: Universidad Austral de Chile.

Corbett, Greville. 2000. *Number*. Cambridge: Cambridge University Press.

Cysouw, Michael. 2003. *The paradigmatic structure of person marking*. Oxford: Oxford University Press.

Daniel, Michael. 2013. Plurality in independent personal pronouns. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. `http://wals.info/chapter/35`.

Elson, Ben. 1960. Sierra Popoluca morphology. *International Journal of American Linguistics* 26. 206–223.

Gaden, Henri. 1909. *Manuel de la langue baguirmienne*. Paris: Ernest Leroux.

Gomes, Antonio Almir Silva. 2013. *Sanapaná, uma língua maskoy: Aspectos gramaticais*: Universidade Estadual de Campinas dissertation.

McCreight, Katherine & Catherine Chvany. 1991. Geometric representation of paradigms in a modular theory of grammar. In Frans Plank (ed.), *Paradigms: The economy of inflection*. 91–112. Berlin: Mouton de Gruyter.

Nussbaum, Loren, William Gage & Daniel Varre. 1970. *Dakar Wolof: a basic course*. Washington, D.C.: Center for Applied Linguistics.

Starke, Michal. 2011. Nanosyntax, part I. Lecture series at GIST, Ghent.

Starke, Michal. 2013. Nanosyntax, part II. Lecture series at CRISSP, Brussels.

Starke, Michal. 2014. Cleaning up the lexicon. *Linguistic Analysis* 39. 245–256.

Stein, P. 1984. *Kreolisch und Französisch*. Tübingen: Niemeyer.

Stewart, William & William Gage. 1970. Notes on Wolof grammar. In Loren Nussbaum, William Gage & Daniel Varre (eds.), *Dakar Wolof: A basic course*. 355–412. Washington, D.C.: Center for Applied Linguistics.

Taraldsen, Tarald. 2012. *ABA and the representation of features in syntax. Talk presented at BCGL 7, Brussels.