Person and Number in Pronouns and the ABA

Guido Vanden Wyngaerd
KU Leuven, CRISSP

1 Introduction

▷ personal pronoun paradigms are multidimensional, in that they involve (at least) the feature dimensions of person and number.
▷ syncretisms may be horizontal (cross-number) and/or vertical (cross-person).

(1)

	sg	pl
1	A	B
2	A	C
3	D	E

	sg	pl
1	A	A
2	B	C
3	D	E

▷ accounting for both horizontal and vertical syncretisms is a challenge for classical nanosyntactic theory

Aims:

▷ to compare two nanosyntactic approaches for deriving syncretisms in multidimensional space:
  ◦ pointers (Caha & Pantcheva 2012)
  ◦ revised Superset Principle
▷ show how the mechanisms of pointers opens the door to the derivation of ABA-patterns.
▷ propose a reformulation of the Superset Principle (RSP) that does not suffer from this defect.
▷ show how empirically, the RSP provides a closer fit with the data.

▷ primary data: Cysouw (2003); subset of the independent pronouns (no agreement markers).

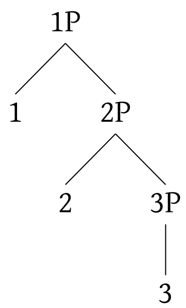
▷ I exclude the inclusive-exclusive dimension from consideration (thorny issue, orthogonal to my concerns).

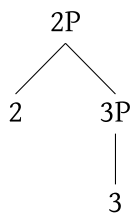
2 Person and number

▷ I assume that the person dimension involves three features ([speaker], [participant], and [person]), which stand in a containment relation (cf. Bejar 2003, Starke 2013)

▷ for convenience, I represent these features as 1, 2, and 3, respectively

▷ the following trees (partially) represent a first, second, and third person pronoun, respectively:

(2) 

(3) 

(4) 

▷ for the number dimension, assume a feature (Num_1) entering into the derivation of singular, and another (Num_2), which is added on top of Num_1 for plural.

▷ some languages form the plural of pronouns with the same morpheme that is used with nouns (or certain noun classes), e.g. Mandarin Chinese (Corbett 2000: 76):

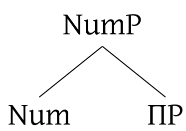
(5)

	sg	pl
1P	wǒ	wǒ-men
2P	nǐ	nǐ-men
3P	tā	tā-men

(6) a. xuésheng
student

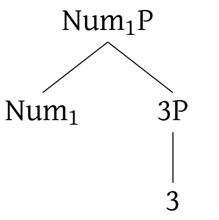
b. xuésheng-men
student-PL

▷ exploiting this analogy, we conclude that the number dimension sits on top of the person dimension complex, as shown in (7):

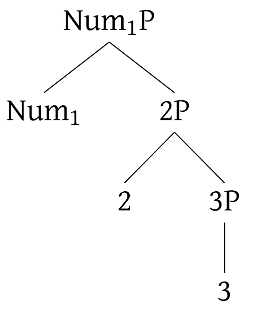
(7) 

▷ assuming that Merge need not construct the full Person fseq before merging Number (i.e. it may leave the fseq incomplete at the top of a dimension), this will yield 6 different syntactic trees (corresponding with the six cells/pronouns in (1)):

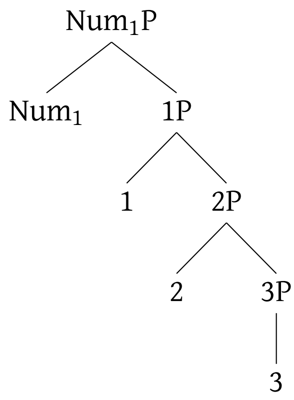
(8)



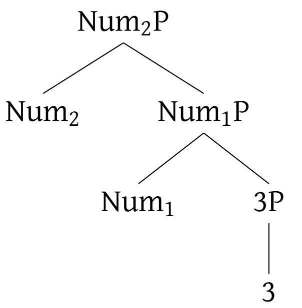
(9)



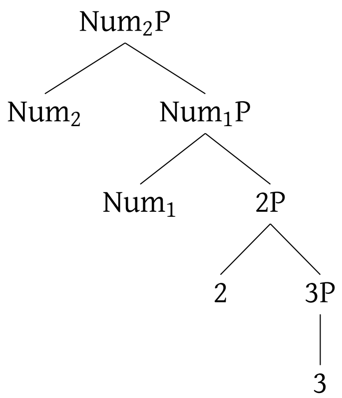
(10)



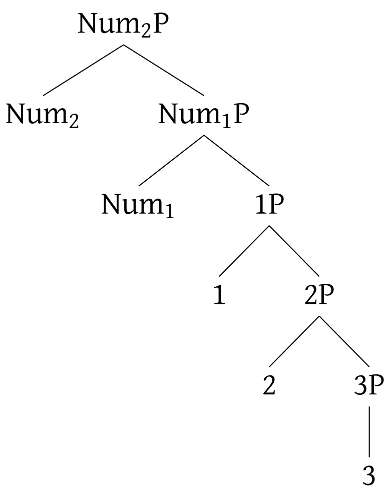
(11)



(12)



(13)



▷ for Mandarin Chinese, assume the following lexical items:

(14)
a. < /tā/, $[_{\text{Num1P}}\ \text{Num}_1\ [_{\text{3P}}\ 3\ ]]$ >
b. < /nǐ/, $[_{\text{Num1P}}\ \text{Num}_1\ [_{\text{2P}}\ 2\ [_{\text{3P}}\ 3\ ]]]$ >
c. < /wǒ/, $[_{\text{Num1P}}\ \text{Num}_1\ [_{\text{1P}}\ 1\ [_{\text{2P}}\ 2\ [_{\text{3P}}\ 3\ ]]]]$ >
d. < /-men/, $[_{\text{Num2P}}\ \text{Num}_2\ ]$ >

▷ the feature composition of the lexical trees in (14abc) corresponds exactly to that of the syntactic trees

▷ to derive the plural pronouns in (5), the complement of Num_2, Num_1P, moves into $\text{SpecNum}_2\text{P}$, after which *-men* spells out Num_2P (spell-out driven movement), deriving the suffixal order.

3 Attested syncretisms

3.1 Types of patterns

▷ vertical (cross-person) ((15)-I)

▷ horizontal (cross-number) ((15)-II)

▷ nonlinear (not exclusively horizontal or vertical) ((15)-III)

(15)

	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

3.2 Horizontal syncretisms

▷ the facts
- 3P: Sinhalese, Sentani, Asmat, SALISH
- specific type: no 3P pronouns, but demonstratives
- 2P (rare): English, Xokleng
- 1P (rare): Marind
- 2P and 3P: Berik, Kuman
- 1P and 3P (rare): Tairora
- all persons: Salt-Yui (3P: demonstratives)

▷ consider Berik:

(16) Berik (New Guinea)

	sg	pl
1P	ai	ne
2P	aame	aame
3P	je	je

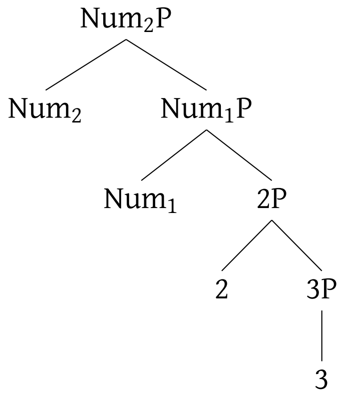
(17) a. < /aame/, $[_{\text{Num2P}}\ \text{Num}_2\ [_{\text{Num1p}}\ \text{Num}_1\ [_{\text{2P}}\ 2\ [_{\text{3P}}\ 3\ ]]]]$ >
b. < /je/, $[_{\text{Num2P}}\ \text{Num}_2\ [_{\text{Num1p}}\ \text{Num}_1\ [_{\text{3P}}\ 3\ ]]]$ >

▷ (17a) can spell out 2P, both singular and plural, by the Superset Principle (Caha 2009):

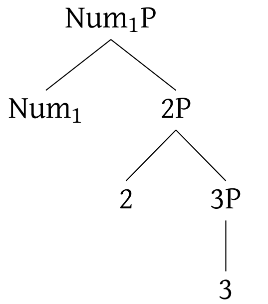
(18) *Superset Principle*
A phonological exponent is inserted into a node if its lexical entry has a (sub-)constituent that is identical to the node.

▷ this condition is met:
- trivially for the 2PL pronoun, as LT and ST are identical;
- for the 2SG pronoun, the Syntactic Tree (ST) (given in (9) above, and repeated here as (20)) is a subtree of the Lexical Tree (LT), as shown below:

(19) Lexical Tree (*aame*)



(20) Syntactic Tree (2SG)



▷ the process by which the LT (19) spells out the ST (20) is sometimes referred to as (metaphorical) 'shrinking': the LT 'shrinks' to become identical to the ST.
▷ the subtree requirement in (18) implies that this shrinking can only happen at the top of the LT.
▷ by the same reasoning, (17b) *je* can spell out 3P, singular and plural.
▷ the other attested patterns of horizontal syncretism work in the same way.

- ▷ the existence of horizontal syncretisms rests on
  - the Superset Principle
  - the possibility to build trees with an incomplete person f_{seq}, i.e. with person features missing at the top of the person sequence

3.3 Vertical syncretisms

- ▷ vertical syncretisms in the singular are extremely rare: Cysouw (2003) finds only one language (out of some 450 listed in the index) showing ABB (Winnebago); Harbour (2015) also lists Sanapaná (Gomes 2013) as having an ABB pattern (both in the singular and the plural).
- ▷ attested patterns in the plural:
  - AAB: many Athabascan languages (e.g. Slave, Chiricahua Apache, Navaho, Kato, Hupa), Awa, Southern Haitian Creole
  - ABB: Nez Perce, Warekena, Wolof (object pronouns), Sanapaná, Mauritian Creole
- ▷ vertical syncretisms cannot be explained under classical nanosyntactic assumptions.
- ▷ consider the AAB pattern in Slave (an Athabascan language, Cysouw 2003: 124):

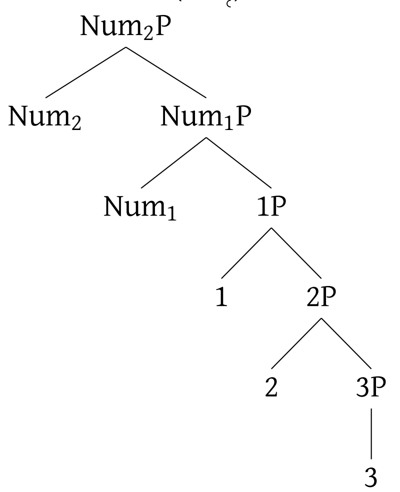
(21)

	sg	pl
1P	sį	naxį
2P	nį	naxį
3P	ʔedį	ʔegedį

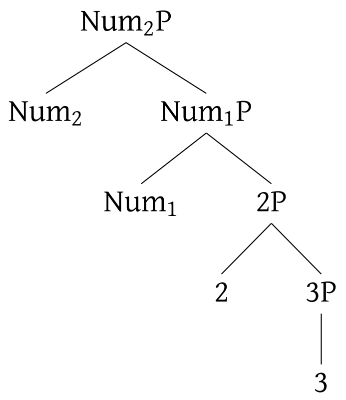
(22) </naxį/, $[_{\text{Num2P}}$ Num$_2$ $[_{\text{Num1p}}$ Num$_1$ $[_{\text{1P}}$ 1 $[_{\text{2P}}$ 2 $[_{\text{3P}}$ 3 $]]]]]$ >

- ▷ (22) can spell out a 1PL pronoun (trivially), but not 2PL one, since a 2PL pronoun is not a subtree of (22) (it lacks the 1P node):

(23) Lexical Tree (*naxị*)



(24) Syntactic Tree (2PL)



- ▷ the ST (24) is not a subtree of the LT (23): the ST is smaller than the LT, but not in the right way: for LT to be able to spell out ST, it would have to shrink in the middle (by deleting the 1P projection).
- ▷ as a result, (23) cannot spell out (24), and the vertical AAB syncretism cannot be derived.
- ▷ for the same reason, the ABB pattern cannot be derived (the lexical item for 2P cannot shrink to 3P)

- ▷ two solutions:
  1. pointers (Caha & Pantcheva 2012): allow trees to shrink a the junctures of the dimensions
  2. a revised *Superset Principle*, which abandons the subtree requirement.

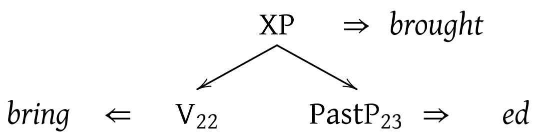
4 Pointers

4.1 Suppletion

- ▷ a pointer is a node in the tree of a lexical item that points to another, existing, lexical item (Starke 2011).

(25)
a. <$_{24}$ /brɔːt/, [$_{XP}$ 22 23]>
b. <$_{22}$ /brɪŋ/, V>
c. <$_{23}$ /ɪd/, PastP>

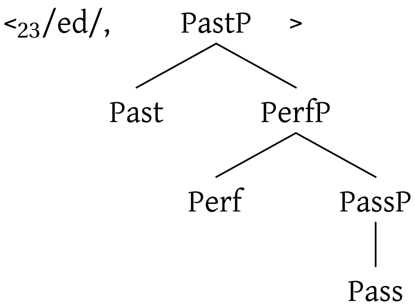
(26)



- ▷ 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 (27), we must conclude that *-ed* has more internal structure, so that instead of (25c), we have (28):

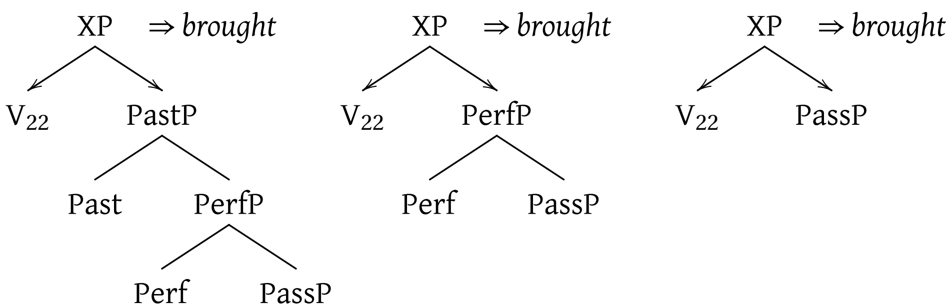
(27)
a. They elected George.
b. They have elected George.
c. George was elected.

(28)



- ▷ the *Superset Principle* ensures that *-ed* may spell out the Simple Past, the Perfect participle, and the Passive participle.
- ▷ the suppletive form *brought* shows the same Past-Perfect-Passive syncretism.
- ▷ this means that in the item with the pointer (26), the item pointed to (28) can shrink to any subtree:

(29)



- ▷ 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.

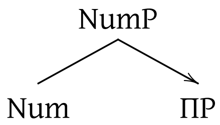
▷ as a result, the lexical item *brought* can spell out three different syntactic trees.

4.2 Multidimensional paradigms

▷ Caha & Pantcheva (2012) propose to exploit the mechanisms of pointers to allow trees to shrink in the middle, and thus account for multidimensional syncretisms.

▷ in this case, lexical items for pronouns could be assumed to contain pointers at the juncture of the dimensions:

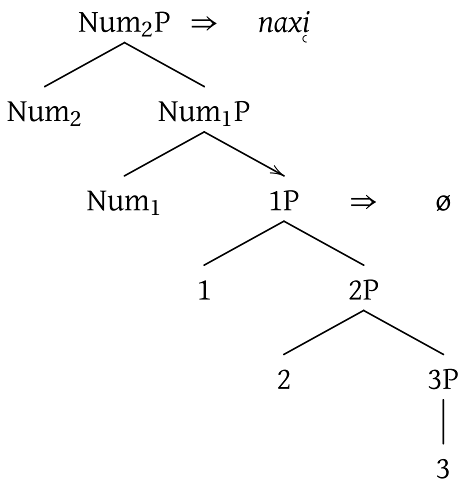
(30)



▷ consider Slave *naxį* again (syncretic for 1PL and 2PL), which we now assume has the lexical entry in (31a):

(31) a. < /naxį/, $[_{\text{Num}_2\text{P}}$ Num_2 $[_{\text{Num}_1\text{P}}$ Num_1 **65**]] >
b. <$_{65}$ /ø/, $[_{1\text{P}}$ 1 $[_{2\text{P}}$ 2 $[_{3\text{P}}$ 3]]] >

(32)



▷ the lexical item (31a) points to lexical item #65, given in (31b).

▷ #65 has a zero phonology and does not occur in isolation, but it is subject to independent cyclic spellout.

▷ this means that #65 can be inserted at a 2P node, by shrinking at the top

▷ #65 will be overwritten at a later stage by *naxį*, when (plural) number has been added.

▷ the vertical syncretism can be derived.
▷ in fact, (31a) can in principle spell out all numbers and all persons.
▷ because of the existence of more specific lexical items in the singular and in the 3P pl, it will lose the competition exactly in those cases, due to the *Elsewhere Principle*.

4.3 Pointers introduce ABA

▷ allowing pointers also allows vertical ABA-patterns (Taraldsen 2012).
▷ consider the hypothetical paradigm in (33), with a vertical ABA, and the hypothetical lexical entries in (34):

(33)

	sg	pl
1P	C	A
2P	D	B
3P	E	A

(34) a. < /A/, $[_{\mathrm{Num_2P}}$ Num$_2$ $[_{\mathrm{Num_1P}}$ Num$_1$ **65**]] >
b. $<_{65}$ /ø/, $[_{1P}$ 1 $[_{2P}$ 2 $[_{3P}$ 3]]] >
c. < /B/, $[_{\mathrm{Num_2P}}$ Num$_2$ $[_{\mathrm{Num_1P}}$ Num$_1$ $[_{2P}$ 2 $[_{3P}$ 3]]]]] >

▷ (34a) and (34b) are like Slave, but (34c) is different from Slave.
▷ the lexical item /A/ (which is identical to that of Slave *naxį̨*) contains a pointer to the lexical item (34b)
▷ /A/ can in principle spell out all persons and numbers.
▷ the ABA-pattern is derived as follows:
  ◦ in 1PL, /A/ is the only candidate and will be inserted.
  ◦ in 2PL, /A/ will lose the competition from /B/ because /B/ is insertable in less environments than /A/:
    – /B/ can shrink at the top ⇒ insertable in 2 cases (2SG and 2PL)
    – /A/ can shrink at the top and in the middle ⇒ insertable in 6 cases (all persons and numbers).
  ◦ in 3PL, /B/ is not insertable because it does not contain a pointer, and only /A/ is a candidate.
▷ if ABA patterns are derivable, this spells bad news for the usability of the syncretism diagnostic to arrange paradigms, and consequently, feature trees.

5 Reformulating the Superset Principle

▷ I propose the following reformulation of the Superset Principle, which gives up the subtree requirement (inspired by Caha 2014):

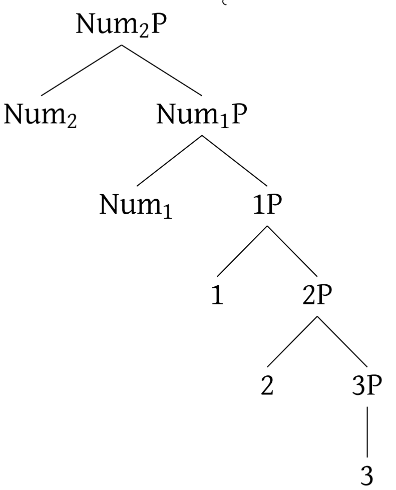
(35) *Revised Superset Principle (RSP)*
A lexical item may spell out a syntactic tree ST iff the features of the lexical item are a superset of the features of the ST.

▷ reconsider the vertical AAB pattern in Slave, and the lexical and syntactic trees for the syncretic pronoun *naxį*:

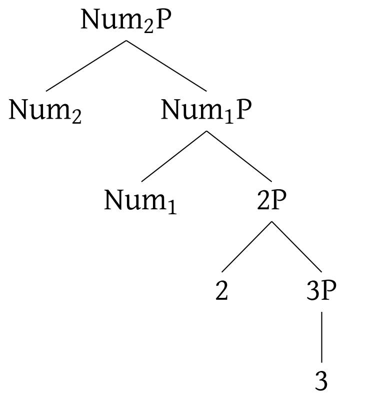
(36)

	sg	pl
1P	sį	naxį
2P	nį	naxį
3P	ʔedį	ʔegedį

(37) Lexical Tree (*naxį*)



(38) Syntactic Tree (2PL)



▷ at the level of the terminals/features, the features of *naxį* (37) are a superset of those of ST (38):
$\{Num_2, Num_1, 1, 2, 3\} \supseteq \{Num_2, Num_1, 2, 3\}$
▷ *naxį* will be able to spell out both 1PL and 2PL by the RSP.
▷ *naxį* could in theory spell out all the cells in the paradigm.
▷ in 3PL, *naxį* will lose the competition from *ʔegedį* because *ʔegedį* is a closer match (similarly in the SG for the SG pronouns).
▷ the vertical AAB syncretism is derived.
▷ the RSP does not allow the derivation of ABA-patterns:

(39)

	sg	pl
1P	C	A
2P	D	B
3P	E	A

▷ if B wins against A in 2PL, it is because B is a closer match for the ST:
  ∘ the syntactic tree for 2PL is smaller than that of 1PL
  ∘ B has less features than A (because A can appear in 1PL)
▷ but then B will also win from A in 3PL, since ST of 3PL has less feature than 2PL
▷ the classical nanosyntactic account for the absence of ABA-syncretisms holds.

Interim Conclusion

▷ both the approach in terms of pointers and the RSP allow the derivation of multidimensional paradigms.
▷ the pointers approach opens the door to the derivation of ABA-patterns, threatening to undercut a cornerstone of the nanosyntactic method.
▷ the RSP abandons the subtree requirement, allows the derivation of both horizontal and vertical syncretisms, and keeps the door on ABA patterns firmly shut.

6 Nonlinear syncretisms

▷ syncretisms which are not exclusively horizontal, and not exclusively vertical either.
▷ they are a special case of multidimensional ones, and therefore require a solution in terms of either pointers or the RSP.
▷ although neither approach provides a perfect fit with the attested data, I show that the RSP provides a closer fit than pointers do.

(40) logically possible ⊃ derivable with pointers ⊃ attested ⊃ derivable with RSP

6.1 L-shaped, contiguous

(41) Usarufa

	sg	pl
1P	ke	ke
2P	e	ke
3P	we	ye

▷ derivable with pointers:
- *ke* is a lexical item containing a pointer; it can spell out all persons and numbers
- *ke* loses the competition to more specific lexical items without pointers (*e, we, ye*)

▷ derivable with the RSP:
- *ke* can spell out all persons and numbers.
- in 2P sg, *ke* loses the competition to the more specific lexical item *e*, which lacks the 1P feature.
- in 2P pl, there is no competition because *e* lacks the plural number node.

▷ the mirror image L-shape of Usarufa is unattested:

(42)

	sg	pl
1P	A	A
2P	A	B
3P	C	D

▷ derivable with pointers (for the same reason as (41)).

▷ underivable with the RSP:
- the A-item can spell out all persons and all numbers.
- the B-item has the same features as A, minus the 1 feature.
- the B-item will therefore win the competiton in both 2SG and 2PL, because of the *Elsewhere Principle*.

6.2 Double L, without ABA

▷ a combination of the Usarufa L-shape with another, interlocking, L in the rest of the paradigm, is unattested:

(43)

	sg	pl
1	A	A
2	B	A
3	B	B

▷ underivable with pointers:
  ◦ both A and B contain pointers.
  ◦ A is maximal and flexible.
  ◦ crucially, B contains a Num_2 feature, as it can spell out 3PL: B is a candidate for 2P and 3P, both SG and PL.
  ◦ A loses out to the more specific B-item in the 3P (unproblematic).
  ◦ A loses to B in the 2SG and 2PL, because B applies to less cases than A.
▷ underivable with the RSP (in contrast to Usarufa above).
  ◦ B has a Num_2-feature (since it occurs in 3PL) and a 2-feature (since it occurs in 2SG), but lacks a 1-feature; therefore it will compete with A in 2P PL, and it will win.
▷ the mirror image of (43) is likewise unattested:

(44)

	sg	pl
1	A	A
2	A	B
3	B	B

▷ underivable with pointers (for the same reason as (43)).
▷ underivable with the RSP.

(Note that both patterns have a vertical syncretism in the singular, which is exceedingly rare independently.)

6.3 Diagonal

▷ one type of diagonal is attested (Suki, Papuan, Cysouw 2003: 121):

(45)

	sg	pl
1P	ne	e
2P	e	de
3P	u	i

▷ diagonal syncretisms contradict spatial accounts of syncretism, which rely on contiguity (e.g. McCreight & Chvany 1991).
▷ derivable with pointers:
  ◦ the lexical tree of the *e*-pronoun is maximal and flexible, i.e. shrinkable at the joint (from 1P to 2P)
  ◦ *e* can express all the persons and numbers
  ◦ it loses the competition against the rigid items for the other persons and numbers

▷ underivable with the RSP:
  ◦ the lexical item *e* can spell out all persons and all numbers (6 cases).
  ◦ in 2P sg, there are two more specific candidates: *ne* and *de*
  ◦ *ne* is predicted to win, because *ne* can spell out 1/2/3P SG (3 cases); *de* can spell out 2/3P SG and 2/3P PL (4 cases), so *ne* is more specific than *de*.
  ◦ the spellout *e* in 2P SG is unexpected.
  ◦ problem for the RSP!
▷ this structure is 'commonly found in the contemporary Aztecan languages' (Cysouw 2003: 121), although from his formulation we cannot be sure if that is true for independent pronouns.
▷ the mirror image of Suki is unattested:

(46)

	sg	pl
1P	A	B
2P	C	A
3P	D	E

▷ derivable with pointers (for the same reason as (45)).
▷ underivable with the RSP
  ◦ both the A and the B-item can spell out all numbers and all persons; the lexical trees of A and B would be identical.
  ◦ there would be a tie between A and B in 1SG, 1PL, and 2PL.
▷ the Suki diagonal plus an L-shape is unattested:

(47)

	sg	pl
1P	B	A
2P	A	C
3P	C	C

▷ underivable with pointers:
  ◦ both A and C contain a pointer, therefore C will win in 2SG.
▷ underivable with the RSP: C is more specific than A and will win in 2SG.

6.4 L-shaped with ABA

(48)

	sg	pl
1	A	A
2	C	B
3	D	A

▷ derivable with pointers:

- the A-item is maximal and flexible; it loses out to the more specific C-B-D items

▷ underivable with the RSP (as are all cases involving ABA).

6.5 Double L, with ABA

(49)

	sg	pl
1	A	A
2	B	B
3	B	A

▷ underivable with pointers:
- there are two competing items, which both contain pointers.
- B will win from A in 3PL since its tree is smaller than the tree of A.

▷ underivable with the RSP.

6.6 Diagonal with ABA

(50)

	sg	pl
1	C	A
2	D	B
3	B	A

▷ underivable with pointers:
- B contains a pointer (to get the diagonal), and will win from A in 3PL.

▷ underivable with the RSP.

Summary

▷ both the approach in terms of pointers and the one in terms of the RSP 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 an attested pattern (the Suki diagonal).

▷ the pointers approach seems too liberal, in that it allows far more syncretisms than are possibly attested (particularly in the diagonal syncretisms, where only one type is attested, but many more allowed by the pointers approach).

7 Further Consequences

▷ the RSP abandons the requirement that there be a subtree relationship between the ST and the LT, and is satisfied with a mere superset relationship between features of the LT and the ST.

▷ as a result, an important piece of motivation is lost for the assumption that lexical items contain trees at all.

▷ the alternative would be that lexical items simply contain unordered sets of features:

(51) a. $< /\alpha/, [_{AP}\ A\ [_{BP}\ B\ [_{CP}\ C\ [_{DP}\ D\ ]]]] >$ (classical nano-LI)
b. $< /\alpha/, \{C, A, D, B\} >$ (RSP-inspired LI)

▷ the classical Superset Principle cannot work with lexical representations like (51b), because of the subtree requirement.

▷ the RSP can work with both types of lexical representations.

▷ some information is lost in a representation like (51b) as compared with the richer kind of representations in (51a).

▷ however, what gets lost is largely redundant information
- the projection of features, i.e. A projects AP, B projects BP, etc.
- the order of these projections

▷ both of these types of information follow from general principles:
- the principle that each head projects
- the universal functional sequence

▷ one area where both approaches to the lexicon make different predictions is in the possibility of the existence of lexical items with recursion in them:

(52) a. $< /\beta/, [_{AP}\ A\ [_{BP}\ B\ [_{CP}\ C\ [_{AP}\ A\ [_{BP}\ B\ [_{CP}\ C\ ]]]]]] >$
b. $< /\beta/, \{A, B, C, A, B, C\} = \{A, B, C\} >$

▷ to the extent that lexical items do not contain recursion, this might be taken as an argument in favour of representations like (52b).

8 Conclusion

- ▷ the analysis of syncretism in multidimensional paradigms requires an extension of classical nanosyntactic theory.
  - ∘ pointers (Caha & Pantcheva 2012)
  - ∘ a *Revised Superset Principle* (RSP)
- ▷ empirically, neither approach provides a perfect fit with the data.
- ▷ conceptually, the RSP is to be preferred as it rules out the derivation of ABA-patterns in principle.
- ▷ assuming the RSP, lexical representations may be assumed to contain unordered sets of features.

References

Bejar, Susana. 2003. *Phi-syntax: A theory of agreement*. Toronto: University of Toronto dissertation.

Caha, Pavel. 2009. *The nanosyntax of case*. Tromsø: University of Tromsø dissertation.

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.

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

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

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

Harbour, Daniel. 2015. *Impossible persons*. MIT Press.

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.

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

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

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