

Ordering paradoxes in a cross-categorial paradigm: on syncretisms with the declarative complementizer*

Bartosz Wiland

Adam Mickiewicz University in Poznań

March 29, 2018

Abstract

This paper shows that incorporating non-definite demonstratives into the same fseq which covers syncretisms with the declarative complementizer, discussed in Baunaz and Lander’s (2018a) work, is necessary to explain syncretic alignment and morphological containment in such paradigms in a systematic way. The paper also resolves an apparent *ABA violation in such a paradigm in Basaá, a language which shows syncretism between the demonstrative and the relativizer to the exclusion of the declarative complementizer.

1 Background: syncretisms with the declarative complementizer

Cross-linguistically attested syncretisms between the nominal declarative complementizer (COMP for short), the demonstrative pronoun (DEM), the interrogative pronoun ‘what’ (WH), and a relative pronoun (REL) suggest that these four categories form a paradigm, understood as an ordered set of related grammatical forms.¹ The syncretic alignment between these categories have led Baunaz and Lander (2017), (2018a), (2018b) to propose that they are syntactically contained as in (1)(where “>” indicates dominance):

- (1) DEM > COMP > REL > WH

*This paper is an abridged version of chapter 5 of Wiland (in prep.) and owes its existence to Michal Starke’s inspiring work on the syntax-lexicon interface, to the editors of this volume, who invited me to contribute, and to an anonymous reviewer who provided feedback on the first draft. Needless to say, I am solely responsible for the statements made in this paper. This work is funded by grant no. 2016/2/B/HS2/00619 (Opus 11) awarded by the National Science Center.

¹I will narrow down the discussion to syncretisms with the nominal complementizer, an equivalent of the English declarative *that*, as opposed to verbal complementizers. The latter are derived from assertive verbs like ‘say’, as for instance the Yoruba *pé*, which is a syncretic form for the verb ‘say’, as in (i), and a declarative complementizer which introduces clauses subordinated under verbs like ‘forget’, as in (ii):

- (i) Olú pé awon ti dé
Olu say they have arrived
‘Olu says they have arrived.’
- (ii) Olú gbàgbé pé Bólá ti jáde
Olu forget COMP Bola PF go.out
‘Olu forgot that Bola has gone out.’ (data from Lawal (1991: 75))

This ordering is motivated by the premise that syncretism anchors structural containment since it only targets contiguous layers of a syntactic structure, a generalization coined in Bobaljik (2007) as the *ABA. The following sample of languages illustrates syncretic alignments consistent upon the order in (1):

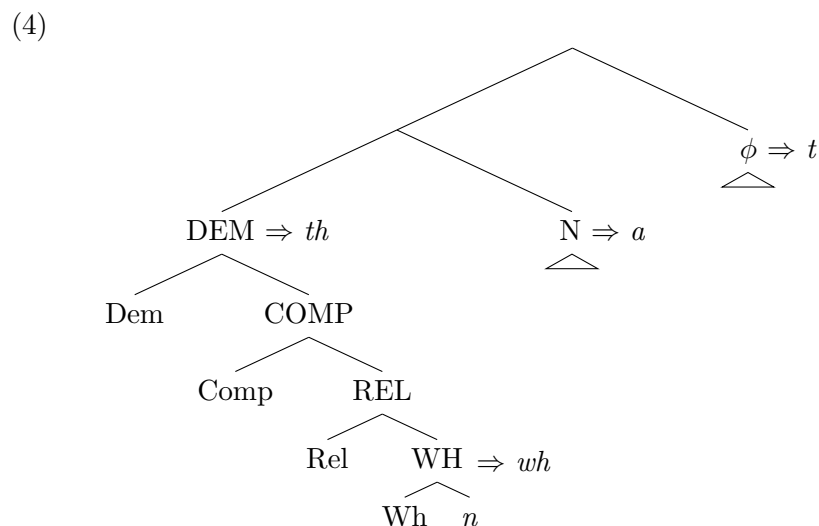
(2)		DEM	COMP	REL	WH
English:		that	that	that	what
Modern Greek:		ekíno	pu	pu	tí
Italian:		quello	che	che	che
Romanian:		acel	că	ce	ce

In nanosyntax, the unattested ABA patterns are a consequence of the spell-out mechanism, which operates in concert with the Superset Principle, defined in Starke (2009) as follows:

- (3) An exponent of a lexical item is inserted into a syntactic node if its lexical entry has a subconstituent which matches that node.

The Superset Principle is amended with the Elsewhere clause, which makes sure that whenever there are lexical items that are in competition for insertion into a syntactic node, it is the exponent of the more specific lexical item (the one with fewer features unspecified in the node) that is inserted.

To briefly illustrate how the Superset Principle rules out the ABA patterns, let us work with the English declarative complementizer, demonstrative, *wh*- and relative pronoun, the categories which in Baunaz and Lander’s analysis have a complex structure as in the tree in (4), whose left branch spells out as the prefix on the nominal stem (marked here as the triangle N), and the right branch spells out the ϕ -agreement features as the suffix.²



Given the lexical entries as in (5-a) and (5-b), *th*- and *wh*- come out as prefixes on *-at* (or rather, strictly following Baunaz and Lander’s description, a bi-morphemic *-a-t*), which spells out an underlying middle and the right branch.

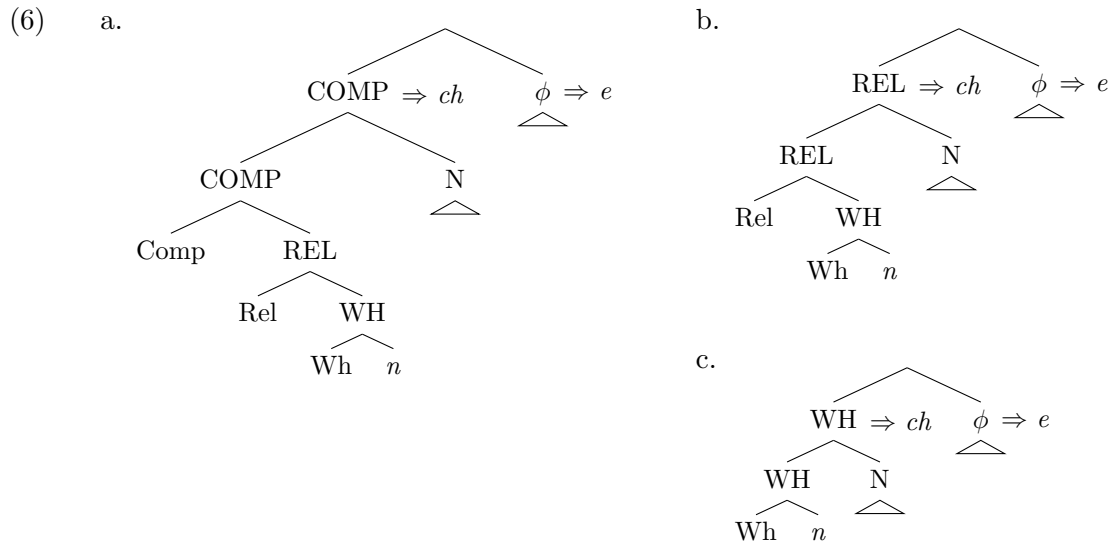
- (5) a. [Dem [Comp [Rel [Wh n]]]] \Leftrightarrow *th*-

²For the sake of concreteness, let us note that the nominal element marked as *n* at the bottom of the left branch of the tree in (4), the stem for the merger of the *Wh* feature, is described as a classifier-like lexical noun in Baunaz and Lander (2018a) and as non-lexical indeterminate noun in Baunaz and Lander (2018b). This issue, however, is largely orthogonal to what follows.

b. [Wh n] \Leftrightarrow *wh-*

The phrasal nodes labelled as DEM, COMP, and REL all spell-out as *th-* as they constitute, respectively, the superset and the subsets of the lexical entry in (5-a). The WH node, also a subset of the entry in (5-a), is spelled out as *wh-* on the strength of the Elsewhere clause, since (5-b) is a more specific match for the WH node than (5-a). The *ABA is, thus, derived in a straightforward way: since the lexical entry for *wh-* is contained within the lexical entry for *th-*, it is impossible for *wh-* to lexicalize a structure bigger than *th-*.

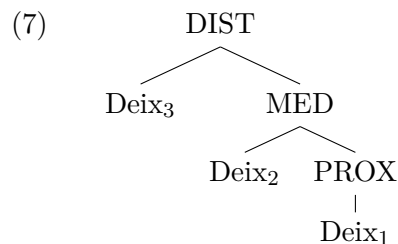
Two remarks are in place before we proceed. First, it is important to note that the labelling used in (4) is a simplified way to illustrate Baunaz and Lander’s analysis, in the sense that a ‘demonstrative pronoun’, a ‘complementizer’, a ‘relativizer’, and a ‘wh-pronoun’ lexicalize both the left branch, the middle, and the right branch of the tree (4) in their analysis, irrespective of morphological complexity of these categories. This is a natural consequence of phrasal spell-out. For instance, in Baunaz and Lander’s system, the *ch-* prefix in the Italian *che* spells out both the left and the middle branch of the representation in (4) as a portmanteau, forming a ‘complementizer’, a ‘relativizer’, and a ‘wh-pronoun’, while *-e* spells out the right ‘ ϕ -agreement’ branch as the suffix, as in:



Thus, the features indicated here as ‘Dem’, ‘Comp’, ‘Rel’, and ‘Wh’ should be understood as subcomponents of the demonstrative, the complementizer, etc., rather than features that solitarily encode the properties of the categories they head. For example, the spatial deictic contrast in English demonstratives *th-*{*is*, *at*} is morphologically realized by the stem, not by the definite *th-* prefix. For this reason, Baunaz and Lander (2018a) describe the ‘Dem’ head in (4) as an instantiation of the definite article, a subcomponent of the demonstrative rather than the source of spatial deixis, an issue that will be taken up in a greater detail shortly. In order to keep the demonstrative pronoun in a language like Basaá, which does not have definiteness marking but shows syncretism with the relativizer in the picture, unless indicated otherwise, I will use the ‘DEM’ label more broadly so that it describes both kinds of demonstrative pronouns. Whenever it will be needed to differentiate between demonstratives with and without definiteness marking, I will refer to them specifically as DEM_{DEF} and DEM_{INDEF}, respectively.

Second, both definite or indefinite DEMs, COMP, REL, and WH should not be necessarily treated as inherently simplex categories beyond their containment. The degree of their internal complexity is irrelevant to the discussion of patterns of syncretism between the categories they form, though. For example, we can follow Lander and Haegeman (2016) and

split the demonstrative into a hierarchy of three distinct deictic features which cumulatively form the proximal (‘close to speaker’), the medial (‘close to hearer’), and the distal (‘remote from speaker and hearer’), as in:



Splitting the deictic component of a demonstrative pronoun in this way is presupposed but it bares no consequences for its ordering with respect to COMP, REL, and WH, so I will continue to refer to it generically as ‘DEM’ for the ease of exposition.

2 Ordering paradoxes with the demonstrative

The DEM=COMP syncretism found in a number of languages, e.g. certain Germanic languages including English, Dutch, and German as in (8), points to the hierarchy ‘DEM > COMP > REL > WH’.

(8)

	DEM	COMP	REL	WH
Dutch:	dat	dat	dat	wat
German:	das	dass	das	was

Some other languages, however, indicate that the order between these categories is different.

One class of challenges to ‘DEM > COMP > REL > WH’ comes from morphological containment of DEM in the structure of the other three categories, which we find in Slavic languages like e.g. Russian or Serbo-Croatian, as in:

(9)

	DEM	COMP	REL	WH
Russian:	to	čto	čto	čto
Serbo-Croatian:	to	što	što	što

The morphological containment of DEM within COMP, REL, and WH is opposite to what we expect if DEM is the biggest of all these categories. This problem is recognized in Baunaz and Lander (2018a), who propose to solve it by eliminating demonstratives without definiteness marking from the sequence so that it applies only to languages with morphologically marked definiteness on demonstratives, that is:

(10) $DEM_{DEF} > COMP > REL > WH$

More precisely, Baunaz and Lander (2018a) argue that only DEM_{DEF} projects as the top layer of the left branch of the tree in (4) and in languages like Russian and Serbo-Croatian DEM_{INDEF} is restricted to the nominal stem, i.e. the middle branch of the tree in (4)(plus the right branch with the ϕ -agreement).

However, such a solution creates a paradox: on the one hand the hierarchy in (10) applies to the categories that always spell-out both the left and the middle branch of the tree in (4)(either synthetically as in English or as a portmanteau morpheme in Italian), on the other hand it is defined only on the basis of the left branch of that tree, excluding the middle branch (putting aside the right branch with the ϕ -agreement, for a moment).

The other challenge comes from (what appears to be) a *ABA violation in Basaá (Bantu, A.43), which shows syncretism between the relative pronoun and the demonstrative to the exclusion of the complementizer, as in:

- (11)
- | | | | | |
|--------|-----|------|-----|-----|
| | DEM | COMP | REL | WH |
| Basaá: | nú | ∅, | nú, | kíí |
| | | lé | lé | |

The status of Basaá as a language with non-definite demonstratives must be qualified. Demonstratives in Basaá can appear in pre- and post-nominal positions. A noun pre-modified by a demonstrative is focused, a noun post-modified by a demonstrative obligatorily takes the augment *í*-prefix, which marks definiteness/specificity (Jenks et al. (2017)), as shown on the example of class 1 demonstrative *nú* (other noun classes are subject to the same constraint):³

- (12)
- | | | | |
|----|--------------------------------|------------|-----|
| a. | í-mut _i | | nú |
| | AUG.1.person | 1.that.DEM | |
| | ‘that person’ | | |
| b. | nú | | mut |
| | 1.that.DEM | 1.person | |
| | ‘THAT person’ (Makasso (2010)) | | |

We can, thus, conclude that while demonstratives in Basaá contextually license an augment prefix which marks definiteness on the head noun, they are themselves not morphologically marked for definiteness.

There are two immediate ways to try to resolve this *ABA problem. One is to go along with the idea that Basaá does not belong to the sequence in (10) since it does not have definiteness morphology (save for the augment marker). This will eliminate the *ABA problem but we will be left with the DEM=REL syncretism which does not belong anywhere in the sequence. On the one hand, placing non-definite demonstratives in an alternative sequence ‘COMP > DEM_{INDEF} > REL > WH’ or ‘COMP > REL > DEM_{INDEF} > WH’ fixes the *ABA violation in Basaá but at the same time it makes the wrong prediction about the morphological containment of DEM_{INDEF} in WH in Russian and Serbo-Croatian. On the other hand, placing non-definite demonstratives at the bottom of ‘COMP > REL > WH > DEM_{INDEF}’ explains the morphological containment in Russian and Serbo-Croatian but it reintroduces the *ABA violation in Basaá in a different place of the sequence by sandwiching the non-syncretic WH between syncretic REL=DEM_{INDEF}.

A different way to go is to assume that since this complementizer is phonologically null, the COMP layer is not projected in Basaá at all. Such an explanation is a little hard to argue since a dialect of Basaá does have a different overt declarative complementizer *lé*, as in (13), which is syncretic with the relativizer, as in (14):

- (13)
- | | | | | | | |
|---|----------|------|-------|----|---------|----------|
| mɛ | í-kâl | lé | Tonye | a | í-kí | yàání |
| I | PRES-say | COMP | Tomye | SM | PRES-go | tomorrow |
| ‘I say that Tonye will go tomorrow.’ (Bassong (2010: ch.3, 30-a)) | | | | | | |

- (14)
- | | | | | | | | | |
|---|----|------|---------|-----|----|-----------|------|--------|
| baúdú | bá | gwě | malět | lé | a | í-kâl | bó | mam |
| students | SM | have | teacher | REL | SM | PRES-tell | them | things |
| ‘The students have a teacher that tells them stories.’ (Bassong (2010: ch.4, 22-b)) | | | | | | | | |

As observed in Bassong (2010), the relativizer *lé* is indeclinable and its distribution in relative

³A syncretic *í*-prefix is reported to function also as a locative marker as in e.g. *í-¹ndáp* ‘at home’ (Makasso (2010: 148), as cited in Jenks et al. (2017)).

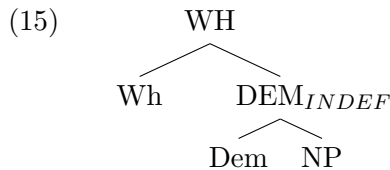
clauses is more restricted than in the case of *nú*, nevertheless, its status as a declarative complementizer must be considered when the nonexistence of the declarative (factive) COMP layer is postulated in order to salvage the hierarchy in (10) for Basaá. An intuitive option would be, thus, to assume that COMP is a skippable layer of structure (an ‘fseq gap’) only on top of the *nú* REL but not on top of the *lé* REL.

In order to incorporate the facts from Slavic and avoid postulating an fseq gap in the *nú*-based paradigm in Basaá, I consider a solution which rests on two independent ingredients: (i) updating the sequence in (10) with DEM_{INDEF} and (ii) showing that the *ABA problem in Basaá is apparent, a conclusion we can draw from inspecting the syntax behind the REL-cell in the Basaá paradigm.

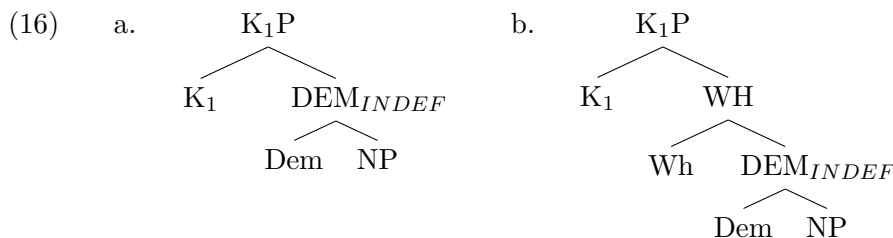
3 Low DEM_{INDEF}

It appears that what constitutes an obstacle in resolving the ordering paradoxes for the sequence in (10) is that it describes the categories realized by both the left and the middle branche of the tree in (4), while the sequence applies only to the properties of the left branch, not the right middle ‘base’ branch (the N triangle). Let us, thus, consider what happens if we relax Baunaz and Lander’s constraint that a demonstrative, a complementizer, a relativizer and a wh-pronoun are always realizations of more than one branch of the tree in (4).

I have argued elsewhere (Wiland (2018)) that the base for the formation of the wh-pronoun ‘what’ in Slavic is the indefinite demonstrative, which constitutes the bottom of a monotonically growing singleton branch, as in:⁴



The nominal ingredient of a demonstrative pronoun – the NP stem that spatial deixis merges – is responsible for the projection of a separate case fseq on its top (marked below as K₁, a stand-in for neuter nominative singular). All these layers are merged in the one and only branch, i.e. they constitute the one and only fseq, as in the structure with a bare DEM_{INDEF} and WH:



⁴More precisely, I’ve argued in Wiland (2018) that the base for the formation of the Polish *co* and Russian *čto* ‘what’ is the medial demonstrative *to*, not the distal *tamto* (Pol) or *eto* (Ru). This makes an interesting prediction about the English *what* as potentially also based on a medial rather than distal *that* (and certainly not on proximal *this* given the non-existent **whis*). An extension of this idea for the English *th-at* and *wh-at*, however, would have to be based on a proviso that the *-at* suffix is either structurally identical in *th-at* and *wh-at* and its realization as in /ð-aet/, /w-ɒt/, /w-ɒt/, or /w-ɔt/ but not as in the expected **/ð-aet/* is a matter of phonology (a possible but unlikely scenario given that /ðae/ is an attested word-initial cluster in Modern English) or that /aet/ in *th-at* and /ɒt/ (and other variants) in *wh-at* are realizations of different subsets of the same fseq. We will return to this issue at the end of section 4.

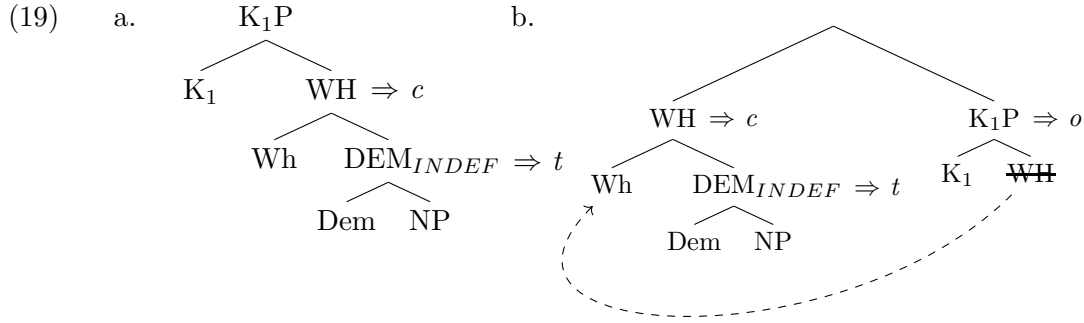
Let us consider how these structures are lexicalized in Polish, a language with bimorphic forms for all four categories, as shown in:

(17)		DEM	COMP	REL	WH
	Polish:	t-o	ż-e	c-o	c-o

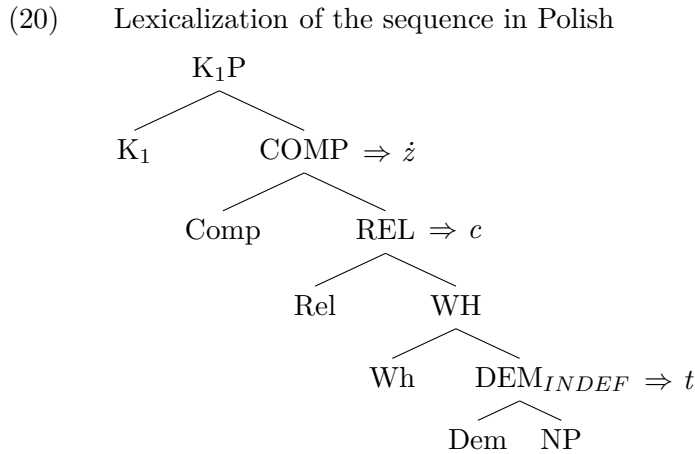
In Polish, the spell-out of the ‘WH > DEM_{INDEF}’ sequence involves a simple over-riding. Given the lexical entries for the exponents of the forms in (17) as in:⁵

- (18) Lexical entries in Polish
- a. [Dem NP] ⇔ t-
 - b. [Rel [Wh [Dem NP]]] ⇔ c-
 - c. [Comp [Rel [Wh [Dem NP]]]] ⇔ ż-
 - d. [K₁] ⇔ -o

the spell-out of the WH over-rides the earlier spell-out of DEM_{INDEF} as in (19-a), and the spell-out of K₁ requires the evacuation movement of its complement, as in (19-b):



There is no need to postulate a second branch (e.g. the N triangle in (4)) if DEM_{INDEF} is already part of WH > DEM_{INDEF}. With the lexical entries in (18), the lexicalization of REL and COMP layers takes place, again, by spelling out the one and only sequence:



Note that the hypothesis that there is a singleton underlying branch for the sequence ‘COMP > REL > WH > DEM_{INDEF}’ does not exclude the possibility that it may have to be reshaped in order to facilitate spell-out. This is a natural consequence of the spell-

⁵The case suffix on the complementizer *ż-e* does not require a separate lexical entry than *-o*. As Baunaz and Lander (2018a) point out, the suffix *-o /o/* shifts into *-e /e/* after a soft consonant *ż- /ʒ/*.

out algorithm in nanosyntax, as summarized in (21), but it does not equal the idea that a reshaped tree is underlyingly multi-branching.

(21) stay → move → subderive

In fact, Starke (2018) advances that the subderivation of the left branch takes place as a last resort operation which facilitates spell-out after ‘stay’ and ‘move’ do not lead to lexical insertion. This is precisely the source of the difference between the pattern we see in Polish and Russian (and Serbo-Croatian), as argued in Wiland (2018). That is, while the shapes of the lexical entries in Polish allow the fseq in (20) to be spelled-out by ‘stay’ (ignoring case), the shape of the lexical entry for the Russian *č* as in (22) requires the formation of the left branch.

(22) Lexical entry in Russian
 [Comp [Rel [Wh Dem]]] ⇔ *č*-

If the lexical entries for the demonstrative *t*- and case *-o* are identical in Polish and Russian, then the lexicalization of WH, REL, and COMP will require the formation of the left branch in Russian, given the entry for *č*-. In contrast to Polish, only the bottom DEM_{INDEF} of the fseq in (20) can be spelled out by ‘stay’ (as *t*-) and movement operations are unable to reshape the tree in (20) in such a way that it matches (the subset or superset of) the entry for *č*-, either. The final available option is to launch a subderivation by providing the last spelled out feature from the mainline, i.e. the ‘Dem’ feature of DEM_{INDEF}, as the basis for the merger of the ‘Wh’ feature, resulting in:

(23)

$$\begin{array}{c} \text{WH} \\ \wedge \\ \text{Wh} \quad \text{Dem} \end{array}$$

Upon the merger of this subderivation with DEM_{INDEF}, the resulting structure comes out as a bimorphemic *č*-*t*- (ignoring, again, the case suffix):

(24)

$$\begin{array}{c} \wedge \\ \text{WH} \Rightarrow \check{c} \quad \text{DEM}_{INDEF} \Rightarrow t \\ \wedge \quad \wedge \\ \text{Wh} \quad \text{Dem} \quad \text{Dem} \quad \text{NP} \end{array}$$

Subsequent mergers forming REL and COMP will extend (what comes out as) the left branch, as discussed in Starke (2018), yielding:

(25) Lexicalization of the sequence in Russian

$$\begin{array}{c} \wedge \\ \text{COMP} \Rightarrow \check{c} \quad \text{DEM}_{INDEF} \Rightarrow t \\ \wedge \quad \wedge \\ \text{Comp} \quad \text{REL} \quad \text{Dem} \quad \text{NP} \\ \wedge \quad \wedge \\ \text{Rel} \quad \text{WH} \\ \wedge \\ \text{Wh} \quad \text{Dem} \end{array}$$

If this is on the right track, then the contrast in the shapes of the lexical items in Polish and Russian informs us that the Polish pattern is more basic, in the sense that the lexicalization of the same fseq is achieved by ‘stay’, while its lexicalization in Russian requires ‘subderive’, the last resort.⁶ We can, thus, conclude that the underlying fseq comprises the non-definite demonstrative at its bottom, as in:

$$(26) \quad \text{COMP} > \text{REL} > \text{WH} > \text{DEM}_{INDEF}$$

Such a sequence is enough to cover the facts from languages like Polish or Russian, but it needs to be updated with definite demonstratives in order to describe languages like English. This issue essentially reduces to the question: “where does definiteness belong to among the other categories?”

4 High DEM_{DEF}

There are at least two scenarios to consider. The first one is a variant of (26) in which definiteness is projected as a separate category at the bottom of the sequence, as in:

$$(27) \quad \text{COMP} > \text{REL} > \text{WH} > \text{DEM}_{DEF} > \text{DEF}$$

Initially, this looks like an attractive option since not only does it suggest that definiteness applies directly to the nominal root, as in (28), but it also reflects the fact that definite markers can be contained in the structure of a demonstrative pronoun (e.g. English *th-at*, or Italian *quel-lo*).

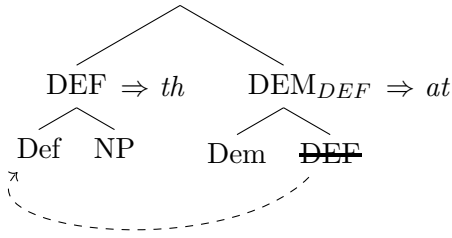
$$(28) \quad \begin{array}{c} \text{DEM}_{DEF} \\ \swarrow \quad \searrow \\ \text{Dem} \quad \text{DEF} \\ \quad \quad \swarrow \quad \searrow \\ \quad \quad \text{Def} \quad \text{NP} \end{array}$$

The idea that definiteness applies to the nominal root parallels with the situation observed with lexical nouns, as e.g. *the car*, where the definite article can appear without demonstrative morphology.

However, extending such a structure into WH, REL, and COMP leads to the *ABA violation: if the English DEF marker *th-* and the medial/distal demonstrative marker *-at* spell out such a structure, the demonstrative *-at* will come out as the suffix, following DEF-movement, as in:

⁶The geometry of the tree in (25) resembles the structure for the Russian *č-t-* as in *č-t-o* in Baunaz and Lander (2018a), where it is based on a multi-branching tree in (4). Note, however, that there are two essential differences between these two representations. The first one is that in Baunaz and Lander’s analysis, the Russian *t-* is an invariant nominal core, a kind of ‘base’ component, while the *t-* in (25) is the medial demonstrative pronoun (modulo the case suffix). The second difference concerns the nature of both representations. In Baunaz and Lander’s analysis, the bi-morphemic *č-t-* realizes the nominal base and the prefix branch of an underlying multi-branching representation in (4). In the alternative in (25), the bi-morphemic *č-t-* is created solely as a result of the spell-out algorithm, hence, there is technically no ‘base’ component or a pre-defined prefix branch; instead, the underlying representation is mono-branching just like it is in Polish (or any other language, for that matter).

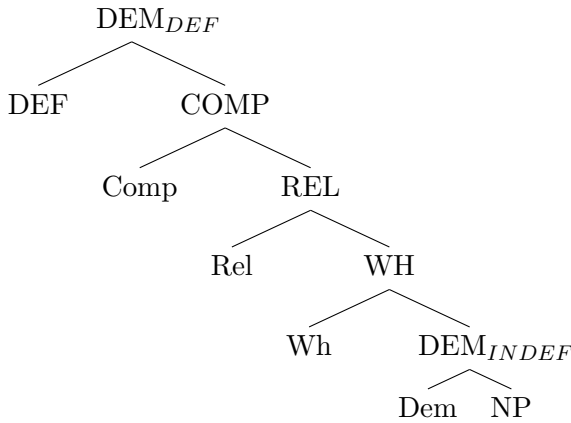
(29)



This appears to be a desired result. However, if the remainder of the sequence is ‘COMP > REL > WH’, then the addition of these layers will result in the *ABA pattern by sandwiching the *wh*- for WH between a lower *th*- for DEF and a higher *th*- for REL and COMP.

In the alternative scenario, definiteness applies to the entire fseq with the nominal root at its bottom, as in:

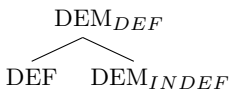
(30) The updated singleton fseq



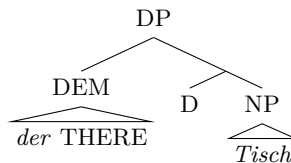
This sequence differs from the one that applies to both Polish and Russian (cf. (20)) only by the top layer and captures the fact that the deictic demonstrative is a stem for the formation of all higher categories.⁷ Its lexicalization in English requires the formation of the complex left branch, as in (31), and with the addition of DEF, it mimics what we see in the lexicalization of this sequence in Russian in (25), modulo the DEF on top.

⁷This option, shown in (i) below without the intermediate WH, REL, COMP layers, is in essence compliant with Leu (2015: ch.2). Leu’s work makes a case for the architecture of the Germanic definite demonstrative which contains the definite article and a ‘proper’ deictic element – an abstract HERE/THERE in Leu’s (p. 15) analysis of German *der Tisch* ‘the table’, as in (ii):

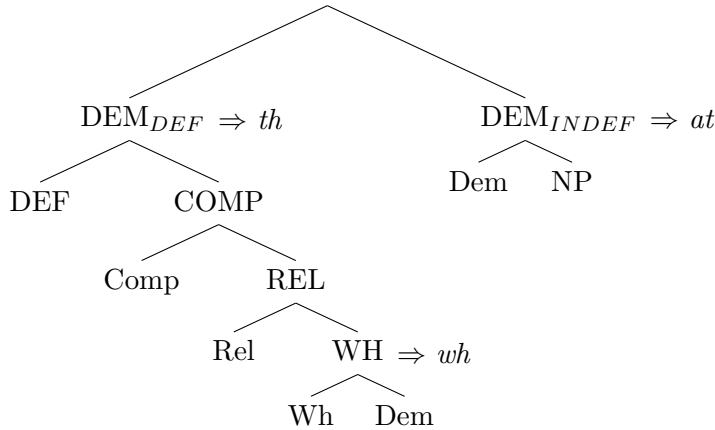
(i)



(ii)



(31) Lexicalization of the English *wh-at* and *th-at*



To sum up, defining the sequence as in (30) leads to the reordering in the paradigms of languages without definiteness marking, which should be represented as in (32).

(32)

	DEM _{DEF}	COMP	REL	WH	DEM _{INDEF}
English:	th-at	th-at	th-at	wh-at	-at
Russian:		č-to	č-to	č-to	to

As already noted in footnote 4, the *-at* morpheme in *th-at* /ð-aet/ and in *wh-at* /w-ɒt/ has different exponents, even across the varieties of English. This contrasts with what we observe in Russian, where *to* is syncretic in all four forms (note that since the Russian *to* is a free form we can easily determine its ‘own’ phonology, which is impossible for the English bound morpheme *-at*). As pointed out by a reviewer, this fact does not appear to result in an ABA pattern, nevertheless, it may suggest that the stem for the formation of the WH, REL, COMP, and DEM_{INDEF} is not constant throughout the English paradigm the way it is in Russian. That is, the English /ɒt/ and /aet/ may reflect a subset-superset relation that is realized by different exponents, a scenario that is in principle possible since the DEM_{INDEF} stem is internally complex.⁸ I leave this issue at this point.

With the sequence that covers both DEM_{DEF} and DEM_{INDEF}, let us return to the second problem, the (apparent) *ABA violation that we find in Basaá, a language with non-definite demonstratives.

5 Resolving the *ABA violation in Basaá

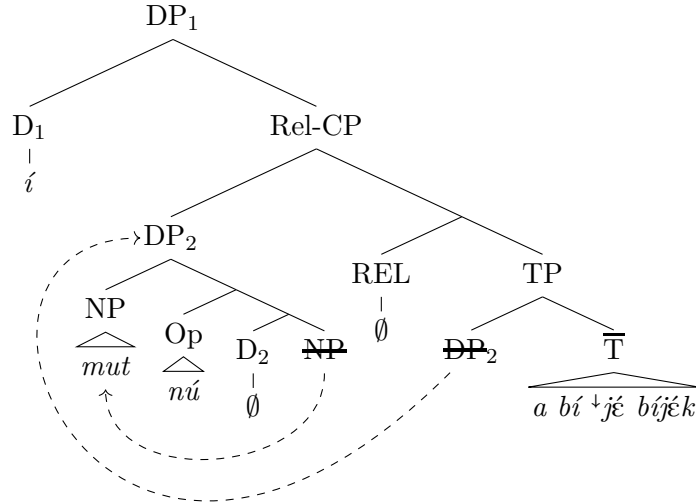
While (30) resolves the morphological containment problem for Slavic, it only pushes the *ABA problem to a different place of the paradigm in Basaá:

(33)

	DEM _{DEF}	COMP	REL	WH	DEM _{INDEF}
Basaá:		∅,	nú,	kíí	nú
		lé	lé		

⁸The complexity of DEM_{INDEF} concerns both the spatial deictic contrast as in Lander and Haegeman’s (2016) decomposition in (7) but also its (pro)nominal component, marked in (30) and elsewhere as the bottom-most NP constituent. In Wiland (2018) I explore a possibility where the Russian and Polish NP *t-* (of *t-o*, as described in (18-a) and (19-a)/(25)) spells out subsets of a nominal sequence specified for THING and PERSON (in the sense of Cysouw (2004)), a scenario more transparently visible in the English forms *wh-at* and *wh-o* rather than in the Russian *č-to* ‘what’ and *k-to* ‘who’ with a syncretic stem *to*.

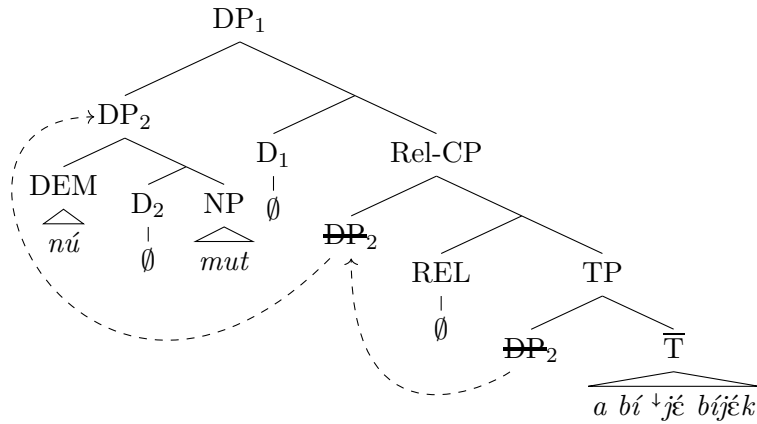
(36)



In such a derivation, the head noun *mut* ‘person’ is first fronted DP₂-internally to a position before the demonstrative “operator” *nú* and in the next step the entire DP₂ is raised to the specifier of the Rel-CP. The selecting head D₁ is lexicalized as the augment *í*-prefix.

Jenks et al. (2017) refer to *nú* in the post-nominal position (36) as the “operator”, which does not receive a focused interpretation in the specifier of Rel-CP. This contrasts with the derivation of the relative clause (34-b) with a pre-nominal *nú* where, as Jenks et al. (2017) argue, *nú* is a genuine demonstrative. In this case, it’s the entire relative DP₂ that moves out of Rel-CP, which results in the demonstrative *nú* receiving a focused interpretation:

(37)



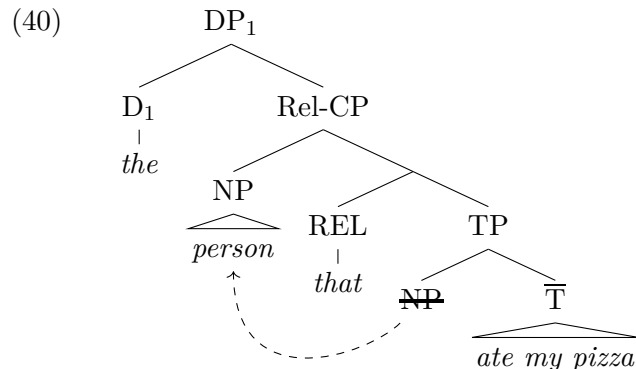
Such an account explains the complementary distribution between demonstratives and (what otherwise would appear to be a distinct category of) relativizers in Basaá, which holds in relative clauses involving a gap – both in subject and object relative clauses and both with pre- and post-nominal demonstratives, with other nominal classes, too, as for instance in:

(38) *í*-maangé_i nú (*nú) [mέ n̄ yí †_i]
 AUG-1.child 1.DEM 1.REL 1sg PR know
 ‘this/that child that I know’

(39) lí lí-wándá_i (*lí†) [†_i lí bí †jé bíjék]
 5.DEM 5-friend 5.REL 5.SBJ P2 eat food
 ‘THAT friend that ate the food’ (Jenks et al. (2017))

If Jenks et al.’s derivation of non-wh-relatives in Basaá is on the right track, then it resolves the *ABA problem in (33) straightforwardly. When we juxtapose the syntax of

non-wh-relatives in Basaá to the syntax of non-wh-relatives in languages like English, we immediately realize that we are not comparing apples to apples and oranges to oranges when we are looking at the REL-cells of both kinds of paradigms. This is so since in relative clauses in languages like English, *that* does not form a constituent with the head noun and, instead, it serves as a genuine relativizer, the head of the relative clause, as in:



In contrast, *nú* in Basaá is a genuine DP₂-internal demonstrative and the relativizer, the head of the relative clause, is null. If so, then the offending paradigm in Basaá should be rewritten as:

(41)

	DEM _{DEF}	COMP	REL	WH	DEM _{INDEF}
Basaá:		∅	∅	kí	nú

6 Summary and conclusion

Cross-categorical syncretisms with the declarative complementizer discussed in Baunaz and Lander (2017), (2018a), (2018b) indicate that the wh-pronoun, the relativizer, the complementizer, and the definite demonstrative pronoun form an fseq. Thus, morphological containment of non-definite demonstrative pronouns in the structure of the wh-pronoun, the relativizer, and the complementizer in languages like Russian poses a problem for such an fseq in that it does not apply uniformly to languages with and without definiteness marking. This can be fixed by inserting non-definite demonstratives at the bottom of this fseq to the effect that the definite demonstrative is a category which syntactically ranges from the non-definite demonstrative, through WH, REL, COMP, and is closed up by a high DEF. This result is possible to achieve if the underlying representation of these categories is simplified to a mono-branch and its partition into multiple morphemes is solely a result of the spell-out mechanism, not the geometry of an underlying representation.

Also, inserting non-definite demonstratives into the fseq that covers the syncretisms with the complementizers provides the first step in resolving an apparent *ABA violation in Basaá, a solution completed by differentiating genuine relativizers (like e.g. in English) from DP-internal demonstratives (or, descriptively speaking, ‘demonstrative operators’) disguised as relativizers in non-wh-relative clauses in Basaá.

References

Bassong, Paul. 2010. The structure of the left periphery in Basaá. Master’s thesis, University of Yaounde I. URL <http://ling.auf.net/lingbuzz/001692>.

- Baunaz, Lena, and Eric Lander. 2017. Syncretisms with the nominal complementizer. *Studia Linguistica* 71:1–34.
- Baunaz, Lena, and Eric Lander. 2018a. Cross-categorical syncretism and the Slavic containment puzzle. In *Balkan syntax and (Universal) principles of Grammar*, ed. Brian Joseph and Iliyana Krapova. Berlin: Mouton de Gruyter.
- Baunaz, Lena, and Eric Lander. 2018b. Deconstructing categories syncretic with the nominal complementizer. *Glossa: a journal of general linguistics* 3 (1):31. URL <http://doi.org/10.5334/gjgl.349>.
- Bobaljik, Jonathan. 2007. On comparative suppletion. URL <http://ling.auf.net/lingbuzz/000443>.
- Carstens, Vicki. 2005. Agree and EPP in Bantu. *Natural Language and Linguistic Theory* 23:219–279.
- Cysouw, Michael. 2004. Interrogative words: An exercise in lexical typology. URL http://cysouw.de/home/manuscripts_files/cysouwQUESTION_handout.pdf, ZAS Berlin.
- Jenks, Peter, Emmanuel-Moselly Makasso, and Larry Hyman. 2017. Accessibility and demonstrative operators in Basaá relative clauses. In *Relative clauses in Cameroonian languages*, ed. Gratien Gualbet Antindogbe and Rebecca Grollemund, 17–46. Berlin: Mouton de Gruyter. URL <http://linguistics.berkeley.edu/~jenks/images/BasaaRelativeClauses.August2015.pdf>.
- Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Lander, Eric, and Liliane Haegeman. 2016. The nanosyntax of spatial deixis. *Studia Linguistica* 71:1–66.
- Lawal, Adenike S. 1991. Yoruba *pé* and *kí*: verbs or complementizers. *Studies in African Linguistics* 22:73–84.
- Leu, Thomas. 2015. *The architecture of determiners*. New York: Oxford University Press.
- Makasso, Emmanuel-Moselly. 2010. Processus de relativisation en Básaá: de la syntaxe á la prosodie. In *Papers from the Workshop on Bantu relative clauses. ZAS Papers in Linguistics 53*, ed. L. Downing, A. Rialland, J.-M. Beltzung, S. Manus, C. Patin, and K. Riedel, 145–158. Berlin: ZAS.
- Ngonyani, Deo. 2001. Evidence for head raising in Kiswahili relative clauses. *Studies in African Linguistics* 30:59–73. URL <https://journals.linguisticsociety.org/elanguage/sal/article/download/1349/1349-2000-1-PB.pdf>.
- Starke, Michal. 2009. Nanosyntax: a short primer to a new approach to language. *Nordlyd* 36:1–6. URL <http://ling.auf.net/lingbuzz/001230>.
- Starke, Michal. 2018. Complex left branches, spellout and prefixes. In *Exploring nanosyntax*, ed. Lena Baunaz, Karen De Clercq, Liliane Haegeman, and Eric Lander. New York: Oxford University Press.
- Wiland, Bartosz. 2018. A note on lexicalizing ‘what’ and ‘who’ in Russian and in Polish. Ms., Adam Mickiewicz University in Poznań.
- Wiland, Bartosz. in prep. *The spell-out algorithm and lexicalization patterns: Slavic verbs and complementizers*. Ms., Adam Mickiewicz University in Poznań.