

Reconciling underspecification with the overlapping decomposition

Tran Truong*

Abstract. Natural language grammar contains pockets of linear morphological contiguity, in which formal identity between non-contiguous categories is either unattested or extremely rare. Also called *ABA domains, the analysis of these etiologically heterogeneous patterns has been a fruitful site of cross-theoretic discourse. Following Caha (2017), surface *ABA distributions may be classified as resulting from one of at least two morphosyntactic decompositions: cumulating and overlapping. The cumulating decomposition characterizes Bobaljik’s (2012) comparative suppletion, and describes domains in which the *final* category of an *ABA domain is the most featurally or representationally complex. In contrast, the overlapping decomposition describes domains in which the *medial* category is the most complex, which Caha further argues is poorly handled by realizational architectures that assume underspecification. This study reconciles underspecification with the overlapping distribution by means of a requirement of contextual contiguity, in which a Vocabulary entry can only realize a complex feature bundle [X, Y] if and only if there exists another entry that realizes either of the simplex features [X] or [Y].

Keywords. morphology; nanosyntax; distributed morphology; contiguity; suppletion; syncretism

1. Introduction. Following Caha (2008, 2013, 2017, *inter multa alia*), *morphological contiguity* is understood in this study to describe linear subsequences within natural language morphosyntax in which absolute syncretisms (Calabrese 2008) and suppletion cannot target non-adjacent subcategories to the exclusion of an intervening category. The best characterized case of a suppletion-sensitive contiguity effect remains Bobaljik’s (2012) comparative suppletion.

	Positive [ADJ]	Comparative [ADJ, CMPR]	Superlative [ADJ, CMPR, SPRL]
AAA	<i>great</i>	<i>great-er</i>	<i>great-est</i>
ABB	<i>good</i>	<i>bett-er</i>	<i>be-st</i>
ABC	<i>bon-us</i>	<i>mel-ior</i>	<i>opt-imus</i>
ABA	<i>good</i>	<i>bett-er</i>	<i>good-est</i>
AAB	<i>good</i>	<i>good-er</i>	<i>be-st</i>

Table 1. Comparative suppletion in English & Latin.

Non-supplementing roots such as *great* maintain formal identity across all three adjectival grades, yielding an AAA pattern. Roots such as *good* co-supplete in the comparative and superlative grades to the exclusion of the positive, yielding an ABB pattern. Although contemporary English does not contain any ABC-supplementing adjectives, this pattern is instantiated by

* Profuse, heartfelt thanks to Karlos Arregi, Itamar Francez, Amy Dahlstrom, Andrew Nevins, Luke Adamson, and my students in the Winter 2020 offering of LING 21000 (Morphology) at the University of Chicago for their help at the earlier stages of this project. Any errors that remain are my own. I am also grateful to Andrew and Chantelle Wells for having kept me fed, watered, and housed during my remote participation at the 95th Annual Meeting of the Linguistic Society of America. Author: Tran Truong (elixir@uchicago.edu).

Latin ‘good’, in which all three roots are distinct. Crucially, the positive and superlative grades cannot co-supplete to the exception of the comparative, an empirical generalization explained by Bobaljik in terms of the *containment hypothesis*, in which the structure of the superlative is built on the structure of the comparative, which is itself built on the structure of the positive. This *cumulating decomposition* is schematized in Table 1 supra by means of incrementally complex toy feature bundles for each grade. Realizational rules sensitive to the superlative environment would necessarily be co-sensitive to the comparative environment.

An additional stipulation is required to explain why *AAB patterns are unattested, even as they appear to be contiguity-preserving at the level of surface allomorphy. Bobaljik proposes an adjacency (or contiguity) condition at the level of the contextual specification of the rule. Specifically, he argues that the trigger for suppletive allomorphy must be adjacent to the root that undergoes allomorphy (Bobaljik 2012:144). This rules out Vocabulary fragments like the following, which would generate *bonus - bonior - optimus*:¹

- (1) *Superlative suppletion without comparative suppletion*
 - a. **Structure of the superlative:** [[[ADJ] CMPR] SPRL]
 - b. **Schema of a rule:** context of insertion ↔ signal / contextual specification
 - c. GOOD ↔ opt- / ___] CMPR] SPRL]
 - d. GOOD ↔ bon-

The containment hypothesis is formalized in (1a), depicting the superlative as containing the comparative. The schema of a realizational rule in Distributed Morphology (Halle & Marantz 1993) appears in (1b), composed of a context of insertion (i.e., a bundle of syntactico-semantic features), a signal (i.e., a phonological exponent), and a contextual specification (i.e., an environment acting as a trigger of allomorphy). By a restriction of *specificational contiguity*, the contextual specification of (1c) is in some sense too complex relative to that of (1d).²

The system in (1c,d) can be made well-formed via the addition of a rule sensitive to the intervening CMPR node:

- (2) GOOD ↔ mel- / ___] CMPR]

More generally, it could be said that a linear contiguity domain does not tolerate a rule that refers to a complex contextual specification [X, Y] unless there is a rule that refers to the simplex contextual specification [X] or [Y].

1.1. BLANSITT’S GENERALIZATION. Caha (2017) describes a linear contiguity domain sensitive to absolute syncretism (or co-lexicalization). Blansitt’s generalization states that in the case subsequence dative-allative-locative, the dative and locative functions cannot be co-lexicalized to the exclusion of the allative. More concretely, DAT refers to the recipient in a ditransitive construction (e.g., *I gave it to Andrew*), ALL refers to the goal of motion (e.g., *I went to the gym*), and LOC refers to ‘the place where’ (e.g., *I am in Utah*). The unattested pattern is a variety of English in which one would say, *I gave it in Andrew*, *I went to the gym*, and *I am in*

¹ As a result of length constraints, this fragment abstracts away from the realization rules for the suffixes.

² Although this study understands the ill-formedness of (1c,d) in terms of a complex trigger of suppletion failing to be licensed by some corresponding simplex trigger, another way of thinking about this is that the target of suppletion (i.e., the root) is too far from a simplex trigger of suppletion (i.e., the SPRL node): GOOD ↔ opt- / ___] SPRL]. Whether the CMPR node belongs to a ill-formed, complex environment *or* whether it acts as an intervener between the target and a simplex environment does not do irreparable damage to the analysis.

Utah, co-lexicalizing the dative and locative functions to the exclusion of the medial allative function.

	Dative [DAT]	Allative [DAT, LOC]	Locative [LOC]
ABC: Basque	<i>-ri</i>	<i>-ra</i>	<i>-an</i>
AAA: Japanese	<i>-ni</i>	<i>-ni</i>	<i>-ni</i>
AAB: Pite Saami	<i>-j</i>	<i>-j</i>	<i>-n</i>
ABB: Dime	<i>-in</i>	<i>-ó</i>	<i>-ó</i>
*ABA: English	<i>in</i>	<i>to</i>	<i>in</i>
Tigrinya	<i>ne</i>	<i>nab</i>	<i>ab</i>

Table 2. Blansitt’s generalization.

From Table 2, theoretically significant differences between this *ABA domain and comparative suppletion are apparent. First, Blansitt’s generalization permits AAB distributions, observable in Pite Saami. Second, and more importantly, the featural architecture proposed by Caha describes the medial category, the allative case, as the most structurally or representationally complex. This characterizes an *overlapping decomposition*, made transparent by data from Tigrinya (as well as Macedonian, Malayalam, and Iatmul), in which the allative is overtly composed of elements from the dative and locative forms.³

This study abstracts away from Caha’s implementation of nanosyntactic overspecification in order to handle these data, focusing instead on his claim that an underspecifying realizational morphology (Stump 2001) would overgenerate *ABA in this domain. It proposes a restriction of *contextual contiguity* targeting the left half of the rule schema in (1b)—complementing the prenominate restriction on *specificational contiguity* targeting the right half—that allows for the reconciliation of underspecification with the overlapping decomposition.

2. Puzzle: Underspecification & the overlapping distribution. Underspecification (i.e., the theoretical claim that phonological exponents need not be fully specified for all syntactico-semantic features) works in tandem with the Pāṇinian Principle⁴ in order to facilitate insertion of phonological content at syntactic terminal nodes. Formally, insertion takes place if the Vocabulary item matches all or a proper subset of the features of a particular context of insertion, and fails if it contains features not present in the morpheme. (That insertion may still take place in the context of ‘extra’ features is characteristic of morphologies that use the Superset Principle, used by Caha in his treatment of Blansitt’s generalization.) Competition between multiple compatible candidates is resolved in favor of the Vocabulary item that matches the most features. Given this massively oversimplified background, one could propose at least two ways to generate the AAB syncretism observed in Pite Sami.⁵

³ The cumulating decomposition is likewise transparent in forms such as Polish *naj-lep-szy* ‘SPRL-good-CMPR’ and perhaps somewhat translucent in English *lattermost*.

⁴ This has also been called the Subset Principle, the Maximal Subset Principle, the Elsewhere Principle, *inter alia*. This study uses the *Pāṇinian Principle* and *elsewhere ordering* interchangeably.

⁵ It should be noted that this a toy implementation that abstracts away from certain details for the sake of brevity. In particular, rule (4b) does not imply an entire absence of any features whatsoever: indeed, this bundle could include more general case-related features. Crucially, (4b) only implies that there are no dative-, allative-, or locative-related features.

(3) *Underspecification with extrinsic ordering*

- a. [DAT] ↔ -j
- b. [LOC] ↔ -n

Rule (3a) generates *-j* as both dative and allative. Rule (3b) generates *-n* as both locative and allative. Given that the allative case is compatible with both *-j* and *-n*, and that both Vocabulary items have the same number of features, it becomes necessary to make a statement of extrinsic ordering, in which rules (or categories) are related to one another in such a way that *-n* is deselected as the allative (e.g., something as simple as DAT > LOC, which amounts to a statement that the dative feature would outrank the locative feature in insertion contexts where they are both compatible). If this is felt to be unparsimonious, another solution is possible.

(4) *Underspecification with elsewhere ordering*

- a. [DAT] ↔ -j
- b. [] ↔ -n

As in (3a), rule (4a) generates *-j* for both dative and allative. An elsewhere form—a form underspecified for any features relevant to this functional subsequence—is generated for the locative. The problem with this approach, according to Caha, is that it overgenerates *ABA, illustrated infra by a hypothetical Vocabulary fragment of an impossible variety of Pite Saami.

(5) *Unattested non-contiguous syncretism of the elsewhere form*

- a. [DAT, LOC] ↔ -j
- b. [] ↔ -n

Rule (5a) generates *-j* for the non-contiguous dative and locative cases. Rule (5b) is an elsewhere form that generates *-n* for the medial allative case. Underspecification appears to be too powerful, in that it predicts that there should be languages in which this pattern of absolute syncretism is active, when none has been observed across genealogically and typologically diverse samples.

3. Solution: Contextual contiguity. Caha is correct that by itself, nothing about underspecification and elsewhere ordering rules out Vocabulary fragments like (5). Something other principle must be at play that keeps learners from positing this system. Just as learners do not tolerate rule configurations that produce *AAB in comparative suppletion, so too do they seem to reject rule configurations that produce *ABA in case co-lexicalization. Specifically, learners do not accept a rule targeting a complex context of insertion (e.g., a bifeatural bundle [X, Y]) in the absence of a rule targeting the constituent simplex context of insertion (i.e., the monofeatural bundles [X] and [Y]). This replicates the dynamic observed in compare suppletion, in which learners reject a rule targeting a complex contextual specification in the absence of a rule targeting the constituent simplex contextual specification.

Intuitively, this suggests that learners disprefer non-incremental jumps in representational complexity. That is, rules like (1c) or (5a) might be difficult to posit in relation to the much simpler elsewhere rules (1d) and (5b), and therefore require the scaffolding of a medially complex rule. Underspecification, the Pāṇinian Principle, and restrictions on contextual and specificational contiguity interact with an antihomophony principle in order to rule out *ABA syncretisms in the linear subsequence that makes up Blansitt's generalization.

(6) *Unattested non-contiguous syncretism via homophony*

- a. [DAT, LOC] ↔ -j
- b. [DAT] ↔ -n
- c. [] ↔ -n

Rule (6a) unproblematically generates *-j* for the allative. Rule (6b) is a dative-specific rule that generates *-n* for the dative, but rule (6c) is an elsewhere rule that homophonously generates *-n* for the locative, yielding the unattested *ABA pattern. Whereas restrictions on contextual and specificational contiguity are conspiring to prohibit the co-occurrence of representations that are *too different* (i.e., representations that are significantly and non-incrementally more complex than others), restrictions on homophony prohibit the co-occurrence of restrictions that are *too similar*.

3.1. PRECEDENT. Adamson (2019) has pursued a similar analysis in his work on English preterites and participles, also a putative *ABA domain.⁶

	Bare [T]	Participle [T, PST, NONFINITE]	Preterite [T, PST]
AAA	<i>try</i>	<i>try</i>	<i>try</i>
ABB	<i>teach</i>	<i>taugh-(t)</i>	<i>taugh-</i>
ABC	<i>drive</i>	<i>driv-(en)</i>	<i>drove</i>
AAB	<i>take</i>	<i>take-(n)</i>	<i>took</i>
*ABA	<i>root</i>	<i>oops</i>	<i>root</i>

Table 3. English preterites & participles.

This *ABA environment shares with Blansitt’s generalization its status as a domain in which the overlapping decomposition operates. More importantly, it shares with comparative suppletion the use of a contiguity restriction that targets the contextual specification. That is, the existence of a rule targeting the bifeatural [PST, NONFINITE] bundle implies the existence of a rule targeting the monofeatural [PST] or [NONFINITE] bundle. The bifeatural rule operating by itself would produce the unattested *ABA configuration.

4. Conclusion. Morphological contiguity phenomena are etiologically and derivationally heterogeneous and amenable to many kinds of theoretical approaches. Specifically, underspecifying realization morphologies need not be incompatible with the overlapping decomposition (i.e., *ABA domains in which the medial category is the most representationally complex), providing that one assumes an additional contiguity restriction at the level of the Vocabulary item. What emerges from cross-domain analysis of natural language morphosyntax is that the grammar contains contiguity requirements that target both the context of insertion (i.e., the feature bundles) as well as the contextual specification (i.e., the nodes acting as distributional constraints on particular allomorphs).

⁶ There is a conversation to be held about how to arrange these categories relative to one another that lies well beyond the scope of this work. Suffice it to say that Adamson makes an excellent case for his particular conceptualization of the data, although in general one could suppose that at least some *ABA configurations are presentational artifacts and not the result of some deeper, abstract linguistic principle.

Domain	Unattested Pattern	Decomposition	Contiguity Restriction
Comparative suppletion	*AAB, *ABA	cumulating	specificational
Blansitt's generalization (case)	*ABA	overlapping	contextual
English verbal root allomorphy	*ABA	overlapping	specificational

Table 4. Heterogeneity in morphological contiguity.

The empirical and derivational diversity of linear contiguity domains is sampled in Table 4. That contiguity restrictions target diverse pockets of grammar (e.g., comparatives, case, and tense) that are characterized by different architectures (e.g., cumulating, overlapping, see also Truong 2020 for discussion of *ABA configurations that seem completely unrelated to any notion of syntactic hierarchical structure) suggest that they are not incidental facts about these domains, distributions, or decompositions. Rather, learners—or Vocabularies, if these are not one and the same—prefer that complex, multifeatureal representations emerge in a stepwise, incremental fashion, in such a way that their presence implies the existence of the corresponding constituent simplex, monofeatureal representation(s). The study ends with a schematization of this claim.

(7) *Contiguity-preserving rule interactions*

- a. **Contextual contiguity:** $[X, Y] \leftrightarrow a$ implies $[X] \leftrightarrow b$ or $[Y] \leftrightarrow c$
- b. **Specificational contiguity:** $[X] \leftrightarrow a / M N$ implies $[X] \leftrightarrow b / M$ or $[X] \leftrightarrow c / N$

References

- Bobaljik, Jonathan. 2012. *Universals in comparative morphology: Suppletion, superlatives, and the structure of [w]ords*. Cambridge, MA: MIT Press.
- Caha, Pavel. 2008. The case hierarchy as functional sequence. In Marc Richards & Andrej L. Malchukov (eds.), *Scales*, 247–276. Leipzig: Institut für Linguistik.
- Caha, Pavel. 2013. Explaining the structure of case paradigms by the mechanisms of nanosyntax. *Nat Lang Linguist Theory* 31(4). 1015–1066. <https://doi.org/10.1007/s11049-013-9206-8>.
- Caha, Pavel. 2017. How (not) to derive a *ABA: The case of Blansitt's generalisation. *Glossa: A Journal of General Linguistics* 2(1). 84. <https://doi.org/10.5334/gjgl.348>.
- Calabrese, Andrea. 2008. On absolute and contextual syncretism: Remarks on the structure of case paradigms and on how to derive them. In Andrew Nevins & Asef Bachrach (eds.) *The bases of inflectional identity*, 156–205. Oxford: Oxford University Press.
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel J. Keyser (eds.), *The view from Building 20*, 111–176. Cambridge, MA: MIT Press.
- Stump, Gregory T. 2001. *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.
- Truong, Tran. 2020. *ABA effects in kinship allomorphy & syncretism. *Proceedings of the Linguistic Society of America* 5(1). 345–353. <https://doi.org/10.3765/plsa.v5i1.4713>.