

Case and Number Suppletion in Pronouns^{*}

Peter W. Smith^{*}, Beata Moskal^{*}, Ting Xu[§], Jungmin Kang[‡]
and Jonathan David Bobaljik[§]

^{*}Goethe-Universität Frankfurt am Main

[‡]Queen's College, City University of New York

[§]University of Connecticut

August 2016

Abstract

Suppletion for case and number in pronominal paradigms shows robust patterns across a large, cross-linguistic survey. These patterns are largely, but not entirely, parallel to patterns described in [Bobaljik \(2012\)](#) for suppletion for adjectival degree. Like adjectival degree suppletion along the dimension positive < comparative < superlative, if some element undergoes suppletion for a category X, that element will also undergo suppletion for any category more marked than X on independently established markedness hierarchies for case and number. We argue that the structural account of adjectival suppletive patterns in [Bobaljik \(2012\)](#) extends to pronominal suppletion, on the assumption that case ([Caha 2009](#)) and number ([Harbour 2011](#)) hierarchies are structurally encoded. In the course of the investigation, we provide evidence against the common view that suppletion obeys a condition of structural ([Bobaljik 2012](#)) and/or linear ([Embick 2010](#)) adjacency (cf. [Merchant 2015](#), [Moskal & Smith 2016](#)), and argue that the full range of facts requires instead a domain-based approach to locality (cf. [Moskal 2015b](#)). In the realm of number, suppletion of pronouns behaves as expected, but a handful of examples for suppletion in nouns show a pattern that is

^{*}Acknowledgements. We are grateful to Daniel Harbour, Martin Haspelmath and two anonymous reviewers for comments on an earlier version of this paper. We would also like to thank Uli Sauerland, Susi Wurmbrand, and audiences at various venues, including those at NELS 45, GLOW 48, Roots IV, SinFonija 9, The 2017 Debrecen Workshop in Pronouns as well as at Bucharest, Cambridge, Concordia, Connecticut, Goethe-Universität Frankfurt, Göttingen, Harvard, the Hebrew University of Jerusalem, Leibniz-Zentrum Allgemeine Sprachwissenschaft (Berlin), Leipzig, Maryland, NYU, Pennsylvania and Vienna. Portions of the work on this paper were carried out under the auspices of a Guggenheim Fellowship to Bobaljik. The support of the foundation is gratefully acknowledged.

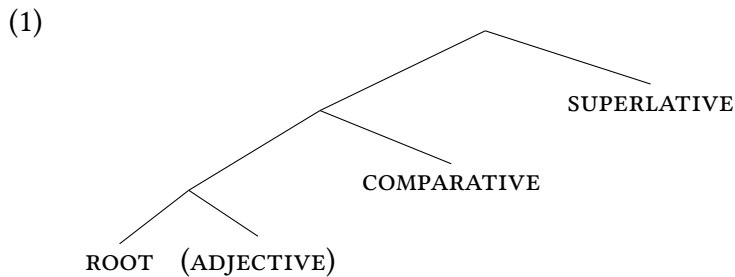
initially unexpected, but which is, however, consistent with the overall view if the Number head is also internally structurally complex. Moreover, variation in suppletive patterns for number converges with independent evidence for variation in the internal complexity and markedness of number across languages.

1 Introduction

Suppletion refers to the phenomenon in which a single lexical item (lexeme) or root morpheme is associated with two phonologically unrelated realisations (exponents).¹ For instance, in the (non-suppletive) paradigm *smart-smarter-smartest* the root remains constant, but in the case of *good-better-best*, *good* ‘changes’ from the root *good* to *be(tt)* in the context of the comparative (and superlative). Suppletion is by definition irregular, however a recent spate of research (Veselinova 2006, Barbiers 2007, Bobaljik 2012, Baerman 2014, Moskal 2015a,b, Bobaljik & Harley 2017) has shown that underneath the surface irregularity lie clear, regular, predictable patterns. For example, in a large cross-linguistic survey of suppletion in comparative and superlative degree formation, Bobaljik (2012) shows that some patterns of suppletion are common, while others are essentially unattested: one finds many examples of an ABB pattern *good-better-best*, in which the comparative and superlative share a suppletive root (B), distinct from the root in the positive (A), but (with some qualifications) no ABA patterns, in which the comparative alone is suppletive: **good-better-goodest*. Bobaljik argues at length that the universality of the patterning requires a structural explanation, and more specifically, that the patterns provide evidence for a largely invariant abstract, internal, hierarchical structure of adjectives. In a nutshell (we elaborate below): the representation of the superlative contains that of the comparative (which in turn contains the root form), as in (1) (a relationship that is morphologically transparent in many languages, though not in English):

Throughout this paper we use the following abbreviations: ABL - Ablative, ACC - Accusative, ADE - Adessive, ALL - Allative, AUG - Augmented, COM - Comitative, COMP - Comparative, DAT - Dative, DL - Dual, EXCL - Exclusive, ERG - Ergative, INCL - Inclusive, INE - Inessive, INSTR - Instrumental, GEN - Genitive, L - low, LOC - Locative, M - mid, NSG - non-singular, NOM - Nominative, OBJ - Objective, PART - Partitive, PCL - Paucal, PER - Perlativ, PL - Plural, POS - Positive, PROX - Proximate, REM - Remote, SG - Singular, SUBJ - Subjective, SPRL - Superlative, UNM - Unmarked, VERS - Versative. We have generally given examples in the orthography used in the source (or a Latin transliteration thereof), and have not tried to standardize to common transcription. We have likewise largely used language names as in the sources cited, without attempting to resolve choices among competing English names.

¹See, for instance, Mel'čuk (1994), Corbett (2005, 2007) for discussions over what should constitute suppletive patterns. Our focus in this paper will be suppletion of pronouns – we take no stand on whether suppletion (root allomorphy) and affixal alternations should be understood in the same terms.



In effect, the absence of the ABA pattern is a consequence of this structural containment – because the superlative contains the comparative, suppletion in the comparative will always preclude the unmarked allomorph in the superlative context. On this account, the unattested patterns do not arise as they cannot be generated in a manner consistent with Universal Grammar.

The goal of the present study is to widen the domain of inquiry to two further complex phenomena to see whether the logic of containment that derives *ABA in Bobaljik’s work holds more generally than just in adjectival suppletion. If the approach in Bobaljik (2012) is correct, we should find that ABA suppletive patterns in particular are excluded (thus unattested) in any sequence of categories that are related via containment. Since the key argument comes from the absence of certain phenomena (*ABA), the domains of inquiry must be large enough to be able to support the claim that the unattested patterns are systematic, and not just accidental gaps. To this end, we investigate suppletion patterns in personal pronouns with respect to morphological case and number. Both areas provide a rich empirical ground for investigating suppletion, and in both domains, we find striking parallels to adjectival suppletion: ABB patterns are robustly attested across language families and are stable over long time periods, while *ABA patterns are unattested. Parity of reasoning with Bobaljik (2012) lends support to theories in which markedness hierarchies for case and number are thus encoded as structural containment of one sort or another. In the course of the investigation, we encounter various differences that point to theoretical refinements: most notably (i) AAB patterns are attested in both case and number suppletion, in contrast to adjectival suppletion (section 3.6), (ii) the interaction of case and number calls into question the general assumption that suppletion is restricted to configurations of structural (Bobaljik 2012) or linear (Embick 2010) adjacency, requiring a somewhat weaker, domain-based relation (cf. Moskal 2015a,b, Moskal & Smith 2016) (section 3.7), and (iii) variation between pronouns and nouns in suppletion for number, along with variation in affix ordering, points to a certain degree of flexibility in structure, which we argue can be modelled via an adaptation of the representation of number in Noyer (1992), Harbour (2008), along with assumptions about the role of markedness in suppletion (section 4.3.1). In the next section, we provide a brief overview of Bobaljik (2012), on which we rely theoretically, and then turn to pronominal suppletion for case and number, respectively.

2 The locality of suppletion: comparatives and superlatives

Bobaljik (2012) conducts a wide cross-linguistic survey into adjectival suppletion in the context of comparative and superlative morphology. Striking in that survey is the finding that not all patterns of suppletion are attested, but rather there are clearly defined patterns with respect to what is, and what is not, possible. The first attested pattern is where there is no suppletion, and the root remains constant as in *smart-smarter-smartest*, referred to as an AAA pattern, since the same root is used in all three degrees. If there is suppletion, then we find two possibilities. First is a pattern where both the comparative and superlative form are suppletive with respect to the positive, but share a common root, an ABB pattern as in the English *good-better-best*. Finally there exist ABC patterns where both the comparative and superlative are suppletive, both with respect to the positive and to each other, seen in the Latin paradigm *bonus-melior-optimus*. The patterns, with selected exemplars, are summarised in Table 1.

Table 1: Attested Patterns of Comparative Suppletion (Bobaljik 2012, 29,56)

		POS	COMP	SPRL	Pattern
a.	English	smart	smart-er	smart-est	AAA
b.	English	good	bett-er	be-st	ABB
c.	Finnish	hyvä	pare-mpi	parha-in	ABB
d.	Latin	bon-us	mel-ior	opt-imus	ABC
e.	Welsh	da	gwell	gor-au	ABC

Strikingly, however there are two suppletion patterns which are not attested. Firstly, there are (virtually) no cases of an ABA pattern.² In other words, there are no adjectival paradigms in which the root is identical in the context of the (regular) adjective and the superlative, to the exclusion of the comparative (Table 2, a.). The second pattern which is unattested is an AAB pattern, where the positive form and the comparative form are constant, whilst the superlative form is suppletive (see Table 2, b.).

Put together, these observations may be stated as follows:

- (2) a. If a root undergoes suppletion in the comparative, it cannot have the positive form in the superlative (roots stay suppletive = *ABA).

²There is one possible counter-example among adjectives of quality from Basque, and a handful of possibly challenging examples from quantifiers: ‘many/much-more-most’. See Bobaljik (2012) for discussion and alternative accounts consistent with the generalisations presented in the main text. In this study we only take into account morphological, or synthetic, constructions and make no predictions for periphrastic constructions.

Table 2: Unattested patterns of adjectival suppletion.

		POS	COMP	SPRL	Pattern
a.	unattested	good	bett-er	good-est	*ABA
b.	unattested	good	good-er	be-st	*AAB

- b. For suppletive morphology to be possible in the superlative, the comparative form must also be suppletive (= *AAB).

In this paper, we aim to build upon these observations and show that (2a) has analogues in other domains of suppletion. The fact that (2b) does not generalise in the same way is anticipated in Bobaljik (2012) and the asymmetry between the two conditions provides further evidence regarding the specific representations involved and the nature of locality in morphosyntactic representations. Before delving into these areas, we first give an overview of how Bobaljik accounts for the attested patterns of suppletion in adjectival suppletion and how the unattested patterns are ruled out.

2.1 Adjectival Suppletion

2.1.1 The Containment Hypothesis

A crucial ingredient for Bobaljik to explain the attested patterns, and rule out the unattested patterns, is the *Containment Hypothesis*:³

- (3) The Containment Hypothesis (from Bobaljik 2012)
The representation of the superlative properly contains that of the comparative.

In other words, the containment hypothesis maintains that the construction in (4) is not a legitimate grammatical object:

- (4) * [[adjective] superlative]

Rather, a superlative construction always has the structure as in (5), crucially containing the comparative.⁴

- (5) [[[adjective] comparative] superlative]

³Bobaljik (2012, Chapter 7) proposes that the Containment Hypothesis is itself a consequence of a deeper condition on the content of functional nodes. Specifically, it is proposed that UG cannot combine the comparative operator MORE and the universal quantifier inherent in the superlative THAN ALL OTHERS into a single functional node (cf. Kayne's (2005, 212) Principle of Decompositionality).

⁴Note that of course not all constructions contain a superlative projection; as such, a comparative is represented as [[adjective] comparative].

The validity of the containment hypothesis is supported by two facts. Firstly, we see in various languages that the containment hypothesis is reflected transparently in the morphology of the forms. This is shown in Table 3 for Czech, Hungarian, and Ubykh, where we can clearly see that the superlative form of the adjective also contains the comparative morpheme.

Table 3: Transparent containment in comparatives and superlatives (Bobaljik 2012, 31)

	POS	COMP	SPRL	Gloss
a. Czech	mlad-ý	mlad-ši	nej- mlad-ši	‘young’
b. Hungarian	nagy	nagy-obb	leg- nagy-obb	‘big’
c. Ubykh	nüs^{oə}	ç’a- nüs^{oə}	a-ç’a- nüs^{oə}	‘pretty’

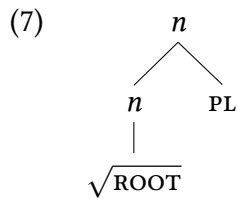
The second piece of evidence suggesting that the containment hypothesis holds cross-linguistically comes from typological universals concerning comparatives and superlatives. Bobaljik formulates the *Synthetic Superlative Generalisation*:

- (6) The Synthetic Superlative Generalisation (SSG)
 No language has morphological superlatives (*X-est*), but only periphrastic comparatives (*more X*).

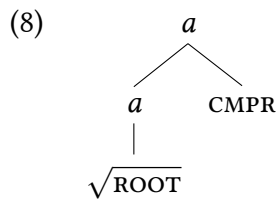
If the Containment Hypothesis is correct, then the SSG (6) effectively follows if the grammar cannot construct a morphological superlative without first constructing a morphological comparative. In other words, a grammar with the resources to construct a morphological superlative must be a grammar that has the resources to construct a morphological comparative.

2.1.2 Distributed Morphology and adjectival suppletion

With the containment hypothesis in tow, Bobaljik shows how the *ABA pattern is excluded within a theoretical framework such as Distributed Morphology (DM, see Halle & Marantz 1993). A key tenet of DM is ‘Late Insertion’: the hypothesis that morphological complexity may be abstract – multi-morphemic words are constructed in the syntax as complexes of abstract terminals (X^0 nodes), which are then subject to post-syntactic rules of exponence (vocabulary insertion, VI). For example, on a DM account, the syntax of a language may represent plural nouns consistently as in (7). The syntax abstracts away from the choice of plural suffix, and it is the post-syntactic, morphological component that determines the overt allomorph of the plural suffix.



In the same vein, a comparative adjective may be represented abstractly as in (8) (we take no stance on the labels of the higher nodes):



Crucially, phonological substance is provided post-syntactically (Late Insertion) and occurs cyclically starting from the most deeply embedded element (Bobaljik 2000). Under this perspective, suppletion may be treated in DM as a special case of contextual allomorphy: more than one VI-rule may compete to spell out a single terminal node. In such a situation, the Elsewhere Principle (Kiparsky 1973), which ensures that more specific VI-rules win out over less specific ones, regulates which exponents are inserted into the syntactic structure during the morphology. In English, there are the following VI-rules for the *good-better-best* paradigm:⁵

- (9)
- a. $\sqrt{\text{GOOD}} \rightarrow \text{be}(\text{tt}) / _] \text{CMPR}]$
 - b. $\sqrt{\text{GOOD}} \rightarrow \text{good} \text{ (elsewhere)}$
 - c. $\text{SPRL} \rightarrow \text{-est}$
 - d. $\text{CMPR} \rightarrow \text{-er}$

In the positive the context for the comparative allomorph (9a) is not met and as such the context-free VI-rule in (9b) applies. In the comparative structure (8) however, the Elsewhere Condition will ensure that the root allomorph (9a) *be(tt)* will be inserted, bleeding the default realization of the same root (9b). This (along with (9d)) correctly derives the form *better*, rather than *gooder*.⁶ Crucially, if we combine the Containment Hypothesis with Late Insertion and the Elsewhere principle, the same VI rules will ensure that in the superlative as well as the comparative, the environment for the more specific VI rule in (9a) is met and as such, (9a) must apply. This result generalises

⁵Note that for the exponents in the VI-rules here, and below, we abstract away from phonological details, and represent them orthographically.

⁶Note that there is no competition or blocking among whole words; the form **gooder* is never derived. See Embick & Marantz (2008) for discussion and comparison with alternatives.

to the more than 100 examples of ABB patterns in Bobaljik (2012): since the superlative contains the comparative, any rule that bleeds the default root allomorph in the comparative will likewise bleed the default root allomorph in the superlative. In other words, *ABA results from the fact that suppletion for the comparative will always entail that the superlative is also suppletive: an ABA pattern cannot be generated by the grammar.⁷

Note that this clearly allows for ABB patterns; it also allows for ABC patterns since a VI-rule can make specific reference to the superlative.⁸ Consider the Latin *bonus-melior-optimus* ‘good’ paradigm, and its corresponding VI-rules (for the roots):

- (10) a. $\sqrt{\text{GOOD}} \rightarrow \text{opt} / \text{ __ }] \text{ CMPR }] \text{ SPRL }]$
 b. $\sqrt{\text{GOOD}} \rightarrow \text{mel} / \text{ __ }] \text{ CMPR }]$
 c. $\sqrt{\text{GOOD}} \rightarrow \text{bon}$ (elsewhere)

The VI-rules in (10b-c) are in relevant respects identical to those in (9a-b). But just as the Elsewhere Condition operates to ensure that the comparative allomorph spreads to the superlative context when there is no more specific exponent, in (10a) there is a more specific root allomorph for the superlative context and thus *opt-* wins out over its competitors.

An additional assumption is needed to exclude AAB patterns. Bobaljik proposes (2012, Chapter 5), in effect, that no grammar may have a rule like (10a) unless it also has a less complex rule like (10b), considering various ways in which this might be derived (see also Bobaljik & Wurmbrand 2013). The existence of AAB patterns turns out to be a point of difference between adjectival degree and pronominal case and number, and we return to this in more detail below.

In sum, the containment hypothesis taken together with the Elsewhere Principle will prevent a root from ‘reverting’ to its original form in the superlative once it has undergone suppletion in the comparative from, and thus ABA patterns are impossible.

3 Case driven suppletion in personal pronouns

Rather than being limited to adjectival suppletion patterns, the same logic should apply to all complex structures where we see evidence for nesting of one structure inside another. In this paper we extend this hypothesis to morphological case and number,

⁷Additional minor rules are needed to ensure that the superlative surfaces as *best* and not **betterest* – see Bobaljik (2012) for discussion. What is relevant for the illustrative point here is that the comparative and superlative share a common root. Since ABC patterns are describable (see immediately below), it is formally possible to mimic a surface ABA pattern, via accidental homophony of A and C. Bobaljik proposes (*ibid.*: 35) to exclude this via a general learning bias against root homophony.

⁸This is somewhat of a simplification especially as regards locality; see Bobaljik (2012) and Moskal & Smith (2016), and section 3.7 below.

with a particular focus on pronouns.⁹ Pronouns are known to show suppletion for both case and number cross-linguistically (Moskal 2015a,b).¹⁰ In addition, pronouns are well described cross-linguistically and we are thus able to construct a large enough sample, such that gaps may be significant. Our general hypothesis is of the form: given a structure in which three (or more) categories stand in a containment relation [[[X] Y] Z], if X suppletes in the context of Y, it will also supplete in the context of Z – that is, there will be no ABA patterns. Working backwards, we may then take the absence of ABA patterns in domains rich with suppletion to constitute evidence of nested structure.

We now turn our attention to the first of the phenomena that are the focus of the present study: morphological case. We will show that the patterns of suppletion that we find with respect to case in personal pronouns show the hallmark *ABA diagnostic of a containment structure for case.

3.1 Why look at case?

It has long been held that the morphological categories of case are subject to a markedness hierarchy, such as in (11) (Blake 1994, 156, Caha 2009, 10):

- (11) NOM < ACC < GEN < DAT < INSTR < COM

More recently, it has been proposed in a number of studies (notably Caha 2009 et seq.) that morphological case is not represented as a simple feature value, NOM, ACC, DAT, etc, but the morphological cases themselves are internally complex, with more complex cases properly containing less complex ones.

There are various strands of evidence that suggest that cases are internally complex. As is the case with degree morphology, there are some languages where case containment is transparently reflected in the morphemes. This is particularly prevalent in locative cases, which are often internally complex, having distinct pieces showing Path and Place (see Radkevich 2010). In addition, oblique cases in many languages are built on a direct case, such as the ergative. The Tabasaran example in (12) shows 4 levels of embedding in the versative case:

- (12) nir -i -q -in -ri
 river -ERG on -ALL -VERS
 ‘towards the bank of the river’ (Tabasaran)

⁹Here and below, we will treat the person formative as the ‘root’ of the pronoun; this is intended loosely - the most deeply embedded morpheme in the pronoun and the one that undergoes suppletion in the cases of interest. We do not intend to take a stand on whether pronouns have roots in some of the technical senses of that term.

¹⁰See in particular Corbett (2005) for an extended argument that alternations in number for pronouns, such as *I* SG ~ *we* PL are genuine instances of suppletion.

Examples of oblique cases transparently containing the accusative can be found in Khanty (also known as Ostyak) pronouns (Table 4, from Nikolaeva 1999, 16) and more systematically in Kalderaš Romani (Table 5, from Boretzky 1994, 31-46). In both tables, the boldfaced morphemes are the case markers contained in the more complex cases.

Table 4: Transparent case containment in Khanty

	NOM	ACC	DAT
1SG	ma	ma:- ne:m	ma:- ne:m -na
3SG	luw	luw- e:l	luw- e:l -na
1PL	muŋ	muŋ- e:w	muŋ- e:w -na

Table 5: Transparent case containment in Kalderaš Romani

	NOM	ACC	DAT	LOC ...
phral	phral- és	phral- és -kə	phral- és -te	‘brother’
phral-(á)	phral- én	phral- én -gə	phral- én -de	‘brothers’
rakl-í	rakl- já	rakl- já -kə	rakl- já -te	‘girl’
rakl-já	rakl- já-n	rakl- já-n -gə	rakl- já-n -de	‘girls’

In addition to the (albeit rare) instances where one case is clearly contained within another, Caha argues that one can formulate implicational universals for whether case is to be expressed morphologically or periphrastically, much in the same way as the SSG of Bobaljik (2012) (see (6) above). Caha proposes the case sequence in (13), purported to hold universally (although he gives a number of important qualifications).

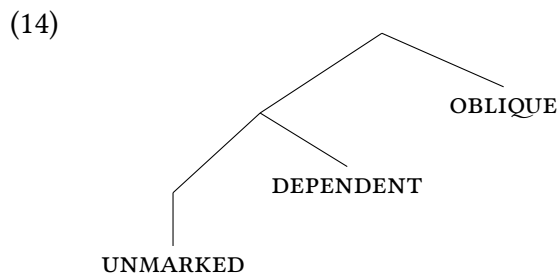
- (13) NOM - ACC - GEN - DAT - INSTR - COM

The final piece of evidence which points towards cases being internally complex is that case syncretisms generally target contiguous regions on the sequence (Caha 2009): Caha contends that a possible syncretism would be one where the accusative, genitive, dative and instrumental are syncretic, however a pattern where the nominative and dative are syncretic to the exclusion of the accusative and genitive is not a possible pattern. By and large, this holds across Caha’s typology of syncretism (but see Harðarson 2016 for a possible counter-example). Abstracting away from the details, Caha shows it is not possible to generate a lexical entry that will target the genitive and the nominative, but not the accusative; in other words, if the genitive and nominative are syncretic, then the accusative must be syncretic with them also.¹¹

¹¹There is a rich tradition dating to work by Roman Jakobson (Jakobson 1936/1971) of using case syncretism to motivate internally complex cases; see for instance McCreight & Chvany (1991), Müller (2004) and Calabrese (2008) among others for somewhat different proposals than Caha’s.

Case, then, turns out to prove an ideal testing ground for a wider application of Bobaljik’s proposal. If the case hierarchy is represented as a containment configuration, then we predict that ABA patterns should be excluded in suppletion.

In order to ensure commensurability among languages, we examined a simplified case hierarchy. Specifically, we considered (i) the unmarked case (the case of canonical intransitive subjects, either nominative or absolutive), (ii) the corresponding dependent case (accusative or ergative), and (iii) a representative oblique case, typically dative. Considerations of markedness, Caha-style syncretic patterning, and what evidence there is for transparent morphological containment point to the structure of case in (14), relative to which we can investigate the attested and unattested patterns of suppletion. We consider the unmarked case the basic form, such that all other cases will at least contain the unmarked case. Dependent case contains the unmarked case only. Oblique cases contain both unmarked and dependent cases.¹² Throughout the study, we set aside the genitive case, as available sources do not consistently distinguish a genitive case (relevant to the case hierarchy) from possessive pronouns (which are not part of the hierarchy).¹³



Our core prediction is relative to this hierarchy: if obliques contain the dependent case (accusative or ergative), which in turn contain the unmarked case (nominative), then an element that shows suppletion for the dependent case will not ‘revert’ to the default form in an oblique. This prediction is systematically borne out for a large sample, to which we now turn.

3.2 Case suppletion

As mentioned above, it is common for pronouns to show suppletion for case (Moskal 2015b), with the Icelandic first and second person pronouns serving as an illustrative

¹²Caha argues that there is a unique, total ordering of containment relations amongst the oblique cases. We do not make that assumption here and allow instead for different obliques to be built from the dependent case, rather than from each other, as suggested by the transparent containment relations in Romani in Table 5, where dative and locative both contain the accusative, but neither contains the other (see also Radkevich 2010, Zompi 2017). We return to this point below in section 3.5.

¹³See also Harðarson (2016) for evidence that the position of the genitive relative to the dative is not universally stable on Caha’s hierarchy. We include genitive and possessive forms in the data in the online appendix.

example (a pattern that is reconstructed for Proto-Indo-European, [Katz 1998](#)). The second person singular is not suppletive, in the sense that it has an invariant person formative *þ-*, followed by a piece that varies for case. The other pronouns in this table, to the extent they are segmentable, show suppletion of the person formative—the first person singular has an *m-* corresponding to 2 SG *þ-* everywhere except the nominative, where *ég* (i.e., [jɛʏ]) starts with a glide *j-*. First and second person plural pattern with the first person singular in this regard, with special nominative forms that share no segments with the corresponding non-nominatives.

Table 6: Pronominal case suppletion in Icelandic.

	NOM	ACC	DAT	
1SG	ég	mig	mér	ABB
2SG	þú	þig	þér	AAA
1PL	við	okkur	okkur	ABB
2PL	þið	ykkur	ykkur	ABB

Since pronouns are ubiquitous and well-documented, it is easy to construct a sizeable cross-linguistic sample. Our initial sample consisted of 160 languages. Of these, roughly half (76) showed no hint of suppletion for case. To the initial sample we added additional cases from the Surrey Suppletion Database ([Brown et al. 2003](#)), and other examples brought to our attention over the course of the investigation. In the end, we considered 179 languages as given in (see [Appendix A.1](#)). Of these, half (90) do not draw more than a two-way case-contrast (after excluding the genitive, as mentioned above). Although some of these languages have suppletion (like English: *I~me*), they are uninformative about the key prediction – they lack ABA patterns, but trivially so, as there is no third category to investigate. Of languages with at least a three-way case contrast, just under half (41) show some suppletion, and are thus informative about the key predictions.¹⁴

We present and discuss the results in the following sections. As the literature makes clear ([Mel’čuk 1994](#), [Corbett 2007](#), [Bye 2007](#), [Haugen & Siddiqi 2013](#)), there is no consensus on where to draw the line between suppletion and other forms of irregularity. Our primary criterion for identifying suppletion was whether or not we could identify an invariant person/number formative for any given pronominal paradigm, as illustrated by Icelandic 2SG = *þ-* above, to be contrasted with 1SG which shows suppletion, alternating between *j-* and *m-*. Yet even with this simple criterion, uncertainties arise. Polish singular pronouns in [Table 7](#) serve to illustrate some of the decision points.

¹⁴The online supplemental material includes the data from all 89 languages with a three-way contrast, since some patterns we exclude as non-suppletive are nevertheless irregular in one way or another, and thus relevant to our interests if other criteria for defining suppletion are used.

(Where multiple forms are in a given cell, the first is a clitic, restricted to unstressed positions. The 3SG *n*-initial variants occur after prepositions.)

Table 7: Selected Polish singular pronouns (Brooks 1975)

	NOM	ACC	DAT	INSTR	LOC	
1SG	ja	mię/mnie	mi/mnie	mną	mnie	ABB...
2SG	ty	cię/ciebie	ci/tobie	tobą	tobie	AAA...
3SGM	on	go/jego/niego	mu/jemu/niemu	nim	nim	ABB

The 1SG forms show the inherited Indo-European ABB pattern, with *j*- in the nominative and *m*- in all other cases. The 2SG alternates between *t*- [t] and *c*- [tɕ]. This alternation is phonological and thus we do not treat it as suppletion (lexical listing of alternants): instead, we posit an underlying /t-/, shared by all forms, which undergoes palatalization before a front vowel or glide (Rubach 1984, Ch. 4). Thus surface *ty - ciebie - tobie* ([ti] - [tɕɛbie] - [tɔbie]) is not suppletive, and is treated as an AAA pattern, and not ABA. Rubach notes, though, that the phonological patterns of palatalization are not trivial and show morphological conditioning to some extent.

Another example which we treat as a morphologically conditioned phonological alternation, rather than as suppletion, comes from Nepali, in Table 8.¹⁵ Of interest are the first person and second person (low honorific) singular, which show diphthongization uniquely before the ergative. With our limited knowledge of Nepali phonology, we have no basis for assuming that the diphthongization here is a general process in Nepali, and grammatical descriptions such as Bal (2007, 359) simply list the forms with a special note. The question of interest, then, is whether the child acquiring Nepali recognises the shared phonology and meaning between *ma* and *mai* (and *tā* and *tai*) and codes this in their grammar, positing a single abstract lexical item and a morphologically restricted phonological rule, or whether they treat this as suppletive, that is, as two separately listed items. We have come down on the side of recognising the clearly identifiable formatives *ma-*, *tā-*, and thus we treat this (and similar alternations) as an AAA pattern with (possibly morphologically conditioned) phonological irregularity, rather than as suppletion.¹⁶ On similar grounds, we recognise third person singular (distal) as having a consistent formative *u(s)-*, which loses the final *s* in the unmarked case.

Returning to Polish, the 3SGM shows another type of consideration. The full pronouns show a pattern shared across Slavic: the nominative pronominal base is *on*,

¹⁵Data from Acharya (1991, 107) and from Sushma Pokharel, personal communication. L and M refer to low and mid honorific grades of the second person.

¹⁶This is of course the same issue that arises with the treatment of “irregularity” more broadly, as famously in the venerable English past tense debate. Our sense of suppletion is narrow, cf. Corbett’s 2007 “maximally irregular” phonology.

Table 8: Selected Nepali singular pronouns.

	NOM	ERG	DAT	LOC	
1SG	ma	mai-le	ma-lāi	ma-mā	AAA
2sGL	tā	taĩ-le	tā-lāi	tā-mā	AAA
2sGM	timi	timi-le	timi-lāi	timi-mā	AAA
3SG	u	us-le	us-lāi	us-mā	AAA

while all non-nominatives are built on a base j -, which surfaces as $(n)j$ - after prepositions. This most likely constitutes an ABB pattern ($on \sim (n)j$ -), if the post-prepositional n is not the same formative as the n in on . On the surface, though, one could imagine seeing the n as a consistent formative (lost after prepositions), thus positing an AAA pattern, or, if one looks only at the weak pronouns (accusative and dative clitics), then one might be tempted to posit an ABC pattern: $on - go - mu$. This last option seems implausible: the clitics (for other gender and numbers as well) are the case endings, with a zero corresponding to the pronominal base. Luckily, while such uncertainty arises in many examples in our database, the important point is that none of the plausible analyses constitutes an ABA pattern. We can thus live with such analytical uncertainty, as far as our primary concern is to test the prediction that ABA will be unattested. In order to provide numbers in the tables below, we have at some points needed to make tricky judgment calls about data points like this, but where a potential ABA pattern is available, as in the Archi 2PL, we have discussed it below. We were conservative in classifying suppletive forms as such, and thus our numbers are if anything, on the low side for suppletive forms. In addition, we have included our full data set as an online appendix, so that the interested reader may pursue this further, and determine whether our results hold up under different analytical assumptions. We return to the general issue in the discussion at the end of the paper.

3.3 Results: Overview

Our results are summarised in Table 9. Consistent with our predictions, there is a fundamental asymmetry between the widely attested Icelandic-like ABB pattern and the virtually unattested ABA pattern. Our original sample turned up but a single possible case of an ABA pattern (more accurately ABBA as it involves multiple obliques), in Archi, which is susceptible to alternative analyses, as discussed in section 3.5 below.

In addition to noting whether patterns are attested or not, we have given numbers from our sample. Note that these numbers are quite conservative, in that they count the number of attested cognate triples of pronouns, not languages (Bobaljik 2012, 40-43). If multiple languages share the same pattern and the elements are cognate, then they are not counted separately. By this count, the common Indo-European 1SG pronouns

Table 9: Pronominal case suppletion: summary

Pattern	Prediction	<i>n</i> attested	Representative languages
AAA	✓	numerous	Lezgian, W. Greenlandic, etc.
ABB	✓	57	Indo-European, Evenki, Khakas, Chuvash, Itelmen
ABC	✓	2	Khinalugh, Albanian
AAB	✓	10	Hunzib, Wardaman
ABA	✗	(1)	Archi 2PL

in Table 10 collectively contribute only a single data point – one of the 57 instances of an ABB pattern in Table 9). The suppletive pattern in Indo-European arose once and has been inherited by the daughter languages, remaining stable over thousands of years.¹⁷

Table 10: Stability of ABB in Indo-European languages.

	NOM	ACC	DAT	Other
German	ich	mich	mir	
Greek	egō	eme	emoi	
Latin	ego	mē	mihi	m-
Lithuanian	àš	manè	mán	man-
Russian	ja	menja	mnje	mn-

Before moving on to the data, we must point out that we restrict our attention to case suppletion in personal pronouns only, and not case suppletion in lexical nouns. This is forced upon us due to an asymmetry between lexical nouns and pronouns discussed in Bybee (1985), Moskal (2015a,b), where it is shown that whilst pronouns frequently supplete for case, barring a handful of examples, lexical nouns essentially never do.¹⁸ The reason for this asymmetry is shown by Moskal to be the result of a difference in structure between functional and lexical material: lexical structure, in contrast to functional structure, contains category-defining nodes, which have the effect of making case information too far away from the lexical root in order to condition suppletion.

¹⁷Although regularized to AAA in, for example, Nepali, as shown above.

¹⁸See Moskal (2015a,b) for a discussion of a limited set of circumstances under which nominal (rather than pronominal) suppletion for case is possible, with analysis of corresponding examples.

Functional structure, by way of contrast, is small enough to allow for case information to condition suppletion on the base of the pronoun. Notably, all of the cases of suppletion in lexical material are shown to involve less structure, which brings case information close enough to the lexical root for suppletion to happen.

3.4 Capturing variation in case suppletion

Our main result, to repeat, is that the predicted patterns are well attested, while the ABA pattern is unattested, with only one dubious potential problem. The AAB pattern turns out to be a point of difference between degree suppletion and pronominal suppletion, and we therefore discuss that pattern in its own right in the next subsection.

As noted above, roughly half of the languages in our initial survey did not show suppletion. Case suppletion is thus not necessary or by any means universal. Select pronouns from Lezgian in Table 11 illustrate a language with rich case distinctions, but no suppletion (AAA).

Table 11: AAA patterns in Lezgian (Haspelmath 1993)

	ABS	ERG	DAT	ADE	INE
1SG	zun	za	zaz	zaw	za
2SG	wun	wuna	waz	waw	wa
1PL	čun	čna	čaz	čaw	ča

Even in languages with suppletion, like Icelandic (6), not all pronouns are suppletive for case, as the contrast between the 1SG and 2SG forms shows. Likewise in English, *I~me* and *she~her* appear to be suppletive, to be contrasted with *they~them*.

Note also that not all AAA patterns are entirely regular. We draw a distinction between suppletive forms, which are built on a phonologically unrelated root, and ‘mere’ irregularity, in which a common root is clearly discernible, despite other irregularities in the form (as discussed for Polish and Nepali above). Along another dimension of irregularity, as Martin Haspelmath points out in comments on an earlier draft of this article, the Lezgian 1SG ergative pronoun lacks an ending that the other ergative pronouns have, and is thus irregular, although it is clearly built on the *z(a)*- base that consistently characterizes 1SG, hence counting as non-suppletive. These patterns were regularly encountered within our survey.

Returning to the data, we now consider ABB forms, where in the dependent and oblique cases, the pronominal base is suppletive relative to the unmarked case, but the suppletive base is consistent. We have illustrated ABB patterns with select Indo-European pronouns in Tables 6, 7, and 10 above. These patterns replicate the basic suppletive patterning seen in degree morphology, and thus receive the same account.

For the Icelandic first person singular, we would set up rules of exponence (vocabulary insertion) such as the following:¹⁹

- (15) a. 1SG → m- / [__] ACC]
 b. 1SG → ég (elsewhere)

Following the logic established for degree suppletion, since the oblique cases, by hypothesis, properly contain the accusative (whether in a Caha-style representation or otherwise – see below), the *m-* allomorph will be used in all the non-nominative cases.²⁰ This approach to suppletion commits us to the view that pronouns are grammatically internally complex; failing to do so (e.g., by listing the various case forms independently as 1SG.DAT → *mér*; 1SG.ACC → *mig*; etc.) would not capture the shared elements among the pronouns, and would not express the patterns of suppletion and syncretism as part of the grammar.

Note that it is not the case that all of Indo-European shares obviously cognate forms. It is not (only) vocabulary items as in (15) that are shared – some of the individual forms have diverged, but the overall pattern remains constant. In addition, some languages have lost a syncretic pattern, while others have innovated – compare the Armenian forms in Table 12, where there is an ABB pattern in 2SG, to those above also from Indo-European, especially the Icelandic data in Table 6 (where 2SG is AAA). It seems then that ABB in Indo-European languages does not stem from an innovation in the set of VI-rules of a single ancestor language, but rather there must be something conspiring to keep ABB wherever there is suppletion.

Table 12: ABB patterns in Armenian (Kozintseva 1995)

	NOM	DAT	ABL	LOC	INSTR
1SG	es	inj	inj(a)nic	inj(a)num	inj(a)nov
2SG	du	k'ez	k'ez(a)nic	k'ez(a)num	kez(a)nov
2PL	duk'	jez	jez(a)nic	jez(a)num	jez(a)nov

¹⁹Above, we have been representing case containment in terms of [[[UNMARKED] DEPENDENT] OBLIQUE]. Since Icelandic has a nominative - accusative case alignment, the case structure for Icelandic is [[[NOMINATIVE] ACCUSATIVE] DATIVE].

²⁰David Adger, Andrea Calabrese, and others have raised the question of whether one could treat the nominative as the marked form, and the non-nominative as the elsewhere case, thus accounting for its wider distribution. This depends on the representation of the unmarked case, e.g., whether the nominative is the absence of case, and thus the larger question of whether rules of suppletion may make reference to the absence of features. For degree morphology, the positive form of the adjective is typically the base for derivational morphology, hence that allomorph should be treated as context-free; but because pronouns do not typically participate in morphological derivation, an analogous argument is hard to construct. We maintain here that the featurally unmarked exponent should be the default, and return to the role of markedness in section 4.3.3.

A selection of Nakh-Daghestanian interrogative pronouns (Kibrik & Kodzasov 1990) illustrates the same point. As shown in Table 13, an ABB pattern remains generally stable, although the actual absolutive forms vary from language to language. Hunzib and Hinuq patterns are given showing that the ABB patterns may regularise, with either the A form or the B form generalising. Yet although the ABB pattern is thus not immutable, what does not arise, as far as we can tell, is an ABA pattern.²¹

Table 13: Suppletion in interrogative pronouns in Nakh-Daghestanian languages

	ABS	ERG	DAT	
Archi	k ^w i	Hi-(li)	Ha-s	ABB
Avar	su	hi-d	hi-bé	ABB
Andi	emi-Ril	He-di-Ril	He-j-Ril	ABB
Bezhta	suk'od	lo-d	lo-l-di	ABB
Hunzib	suk'u	suk'u-l	suk'-u	AAA
Hinuq	lu	luj	luz	AAA

Itelmen (Khairjuzvo dialect forms are given in Table 14) may also provide an example – the 1SG pronoun is regular, but the 2SG pronoun lies in the grey area between suppletion and irregularity. The root alternates between unmarked *kəzza* and the root *kn-* in all other cases. The historical phonology of this alternation can be reconstructed on the basis of comparative Chukotko-Kamchatkan evidence, but synchronically, it is difficult to see a motivation for treating this as anything other than suppletion, and hence, an ABB pattern.

Table 14: ABB (2SG) in Itelmen (Bobaljik, Field Notes)

	UNM	LOC	DAT	ABL
1SG	kəmma	kəmma-nk	kəmm-ank-e	kəmma-n-xʔal
2SG	kəzza	kni-nk	kn-ank-e	kna-n-xʔal

Finally, we note that ABC patterns, as with adjectival suppletion, are exceedingly rare, but appear to be attested. A selection of Nakh-Daghestanian 1SG pronouns are given in Table 15 (from Kibrik & Kodzasov 1990). While the majority of these, like Lezgian above, show a constant formative across the cases (thus constitute AAA patterns), the Khinalugh pronoun appears to be deviant, showing an ABC pattern.²²

²¹The Andi form is an ABB pattern: *emi-He* is the *wh*-root; *-Ril* is a suffix that distinguishes, according to the description, ‘known’ from ‘unknown’ *wh*-words.

²²The Albanian third person singular pronoun may also be an ABC pattern but is less clear; see n. 95 below.

Table 15: 1SG pronouns in some Nakh-Daghestanian languages.

	ABS	ERG	DAT	
Khinalugh	zi	jä	as(ir)	ABC
Chamalal	di:	de:	di-la	AAA
Rutul	zi	za-d	za-s	AAA
Tabasaran	izú	izu	izu-s	AAA

To close this section, we note one further paradigm that shows some of the difficulties in characterising pronominal suppletion in a large sample of languages, not all of which we are familiar with. The pronouns of Murle are given in Table 16. How should these be classified?

Table 16: Murle pronouns ([Arensen 1982](#))

	NOM	ACC	DAT
1SG	naana	aneeta	ɲaatan
2SG	niina	ineeta	ɲaatun
3SG	niini	nɔnnɔ	ɲaatin
1PL	naaga	ageeta	ɲaatinaaŋ
2PL	niiga	igeeta	ɲaatiŋooŋ
3PL	niigi	ɲɔɔgɔ	ɲaatineeŋ

Looking at any of the personal pronouns in isolation, one may be tempted to see these as an ABC pattern, but looking at the paradigm as a whole, it is clear that there is a great deal of systematicity. The local (1,2) pronouns are built on the pattern: NOM = *n-X-a*, ACC = *X-eeta*, DAT = *ɲaat(in)-X*; where the X is a constant person/number formative, varying only slightly in vowel length (and vowel quality in the dative): 1SG = *(a)an*, 2SG = *(i)in~un*, etc. While the fine detail of the segmentation may be tricky, it would seem unwise to build a case for an ABC pattern on examples of this sort, and we tentatively treat the Murle pronouns as non-suppletive, noting that nothing of consequence for the theory hinges on this choice. Either they are to be treated as AAA or ABC - the choice affects the numbers in Table 9 but the main prediction that ABA should be unattested is not affected by this analytic uncertainty.

3.5 Apparent ABA - Archi 2PL

There is only one potential candidate for an ABA pattern in our sample: the 2PL form in Archi. Table 17 illustrates the Archi system for first and second person pronouns.

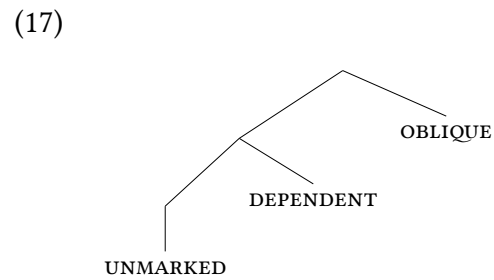
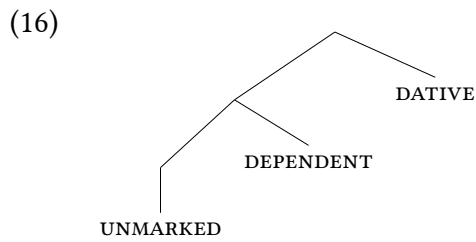
Before turning to the ABA form, the Archi dative form in Table 17 deserves additional comment. In a typologically rare pattern (though to some degree attested in related languages), certain Archi pronouns show agreement in gender/noun class with a clausemate absolutive argument. This is indicated as \varnothing (for gender marker) in Table 17, a prefix that surfaces as $\{w-, d-, b- \text{ or } \emptyset\}$. See Polinsky et al. (In press) for discussion of this phenomenon.

Table 17: Pronouns of Archi.

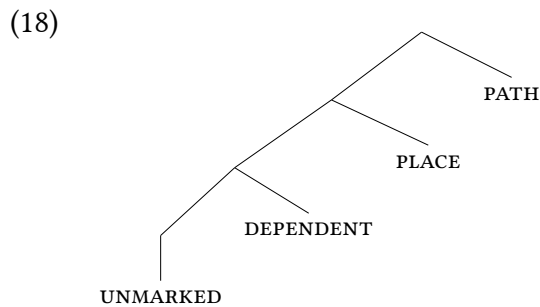
	ABS	ERG	DAT	OBL
‘who’	k ^w i	Hi-	Ha-	–
1SG	zon	za-ri	\varnothing -ez	za-
1PL.EXCL	nen	nen	\varnothing -el	la-
1PL.INCL	nen	nen- \varnothing	\varnothing -el-a- \varnothing -u	la-
2SG	un	un	wa-s	wa-
2PL	ž ^w en	ž ^w en	wež	ž ^w a-

The potential problem for us comes from the final row of the table, in the 2PL form. The other rows all show patterns consistent with our predictions: The interrogative pronoun is clearly an ABB pattern, and the 1SG form has a consistent formative *z* throughout all cases. The 1PL and 2SG forms are apparently AABB patterns, at least according to the column alignment here (which we will revise immediately below), thus counting as either AAB or ABB depending in part on the treatment of syncretism (see below). But the 2PL form appears to show an elsewhere base \check{z}^w -, with apparent suppletion in the dative (*wež*), before apparently reverting back to the non-suppletive root for more complex oblique cases. However, it is not clear that this constitutes an ABA pattern of the type that should be predicted not to exist under our proposals.

The worrisome sequence arises only if the oblique cases properly contain the dative, as on Caha’s approach. In the discussion of Kalderaš Romani, above, we noted that transparent containment relations suggest instead that the various obliques, including the dative, are each separately built from the dependent case, while the obliques do not contain one another. That is, we assumed above that the cases beyond the unmarked and dependent constitute a partial, rather than a total, order, including (16) and separately, (17):



It seems to be quite generally true cross-linguistically (Radkevich 2010) that spatial cases conform to an abstract structure such as that in (18), where each of the nodes PLACE and PATH may correspond to a range of values, but none of which are systematically contained inside the another.²³



An illustration that this is on the right track comes from Lezgian (Haspelmath 1993), who gives the paradigm for spatial cases, illustrated by *sew* ‘the bear’ shown in Table 18. The spatial cases all transparently conform to (18). All contain the ergative morpheme *re*. Beyond that, the cases may be grouped into series, with the first element *w*, *q^h*, *k* or *l* denoting position, and the second denoting motion to, from, or location at. But it seems implausible to posit that all 12 local cases form a unique sequence, where, for example, either the subrelative includes the superdirective or vice versa.

Given such evidence, it seems reasonable to treat with some suspicion the notion that Archi 2PL forms constitute an ABBA pattern. Instead, it seems plausible to assume that there are two distinct containment patterns. The Archi dative corresponds to (16) and shows an AAB pattern, while the other obliques correspond to (17) and show an AAA pattern. On this view, independently motivated by the transparent embedding patterns within Daghestanian, no ABA pattern is in evidence, as predicted.²⁴

²³Radkevich’s structure is more articulated than the one given here. In addition, she argues that patterns of portmanteau morphology suggest that PLACE and PATH (and their dependents) form a (surface) constituent, to the exclusion of the dependent case node. See Pantcheva (2011) for an approach which posits a total order among the local cases.

²⁴Even if we set aside the possibility of a partial, rather than a total, order among the oblique cases, it may be possible to analyse the final *z* in the dative as constituting the same formative as the initial

Table 18: The spatial case paradigm of Lezgian

Case		Translation
ABSOLUTIVE	sew	‘the bear’
ERGATIVE	sew-re	‘the bear’
DATIVE	sew-re-z	‘to the bear’
ADESSIVE	sew-re-w	‘at the bear’
ADELATIVE	sew-re-w-aj	‘from the bear’
ADDIRECTIVE	sew-re-w-di	‘toward the bear’
POSTESSIVE	sew-re-q ^h	‘behind the bear’
POSTELATIVE	sew-re-q ^h -aj	‘from behind the bear’
POSTDIRECTIVE	sew-re-q ^h -di	‘to behind the bear’
SUBESSIVE	sew-re-k	‘under the bear’
SUBELATIVE	sew-re-k-aj	‘from under the bear’
SUBDIRECTIVE	sew-re-k-di	‘to under the bear’
SUPERESSIVE	sew-re-l	‘on the bear’
SUPERRELATIVE	sew-re-l-aj	‘off the bear’
SUPERDIRECTIVE	sew-re-ldi	‘onto the bear’

Similar considerations may apply to Nen (Evans 2015, 553), another apparent AAB...A example not in our own survey but brought to our attention by Greville Corbett (personal communication). The Nen pronominal paradigms for seven of the twelve cases identified by Evans are given in Table 19:

No issues arise for personal pronouns within the limited sequence unmarked (absolute) – dependent (ergative) – basic (zero-suffixed) oblique, and we note both AAB (1st person) and ABB (3rd person) suppletive patterns relative to this sequence. As Evans notes, most of the oblique cases (including the 5 not shown) are built on the oblique stem, with further containment relations – 4 obliques, like the allative and ab-

z^w - in the other forms, and thus an AAA pattern. Nina Radkevich calls our attention to Alekseev (1985, 70-75), who analyzes both the genitive and dative as arising (historically) from metathesis of z and w , plus a vowel change, and finds evidence for the components of this analysis in related languages. This analysis may be supported by analogy to the Archi 1sg forms, which show a similar pattern, including devoicing in the genitive (1b) (see Moskal 2013, Alekseev 1985):

(1) 2ND PLURAL

2ND PLURAL			1ST SINGULAR			
a.	ABS	\check{z}^w	-en	b.	z	-on
	ERG	\check{z}^w	-en		za	-ri
	GEN	w-	iš	∅-	is	
	DAT	w-	ež	∅-	ez	
	OBL		\check{z}^w a-		za-	

Table 19: Nen pronouns in selected cases

	ABS	ERG	OBL	DAT2	ALL	ABL	COM	PER
1SG	ynd	ynd	ta	tapap	tapapt	tapapngama	yndba	yndma
1NSG	ynd	yndbem	tbe	tbepap	tbepapt	tbepapngama	yndba	yndma
2SG	bm	bm	be	bepap	bepapt	bepapngama	bmba	bmma
2NSG	bm	bmbem	bbe	bbepap	bbepapt	bbepapngama	bmba	bmma
3SG	bä	ymam	ya	yapap	yapapt	yapapngama	ymaba	bäma
3NSG	bä	ymam	ybe	ybepap	ybepapt	ybepapngama	ymabeba	bäma

lative, properly contain the dative 2, both in terms of root and affixes. The outliers of interest are the comitative, which is built on the ergative, and the perlative, apparently built directly on the absolutive. That the comitative is built on the ergative is shown by the suppletion in the third person, and also, as Evans notes, by the curious fact that the singular interrogative pronoun root *ebe* is in the ergative syncretic with the third person singular pronoun, and this odd syncretism is inherited by the comitative. The comitative and the basic oblique thus further illustrate the non-total ordering of oblique cases, as in (16) and (17). Thus what stands as the potential problem for us is the third person perlative case, namely the sequence *bä – ymam – bäma*. The third person perlative shares a base with the absolutive, not used for the ergative (or any of the other 9 cases). As Evans notes, the perlative appears to be built directly on the absolutive. This is surprising from our perspective, since it runs against the view that all obliques should obey at least the partial ordering in (14), embedding at least the dependent case. Evans notes that the perlative bears a special relation to the absolutive in another way: a typological oddity of Nen is that the absolutive fails to mark number distinctions (for pronouns and nouns alike) that are marked in the other cases (e.g., by non-singular *-be-* in Table 19) – this number-neutrality, as can be seen in the final column of Table (19) is inherited by the perlative. We leave the analysis of the Nen third person perlative unresolved here, although we note that in addition to the issues above, a further issue that may come into play is the fuzzy boundary between suffix and postposition.²⁵ Bobaljik (2012) argues that the trigger for suppletion must be (effectively) within the same morphological word as the target – a suffix should be a possible trigger for suppletion, but if the perlative could be analysed as a postposition, selecting a caseless complement, no suppletion would be predicted, or for that matter, possible.²⁶

²⁵See, for example, Trommer (2008) and Spencer & Stump (2013) for opposing views on the treatment of oblique case suffixes in Hungarian as case affixes or as phonologically dependent postpositions.

²⁶By contrast, the comitative pattern would not be problematic even if the comitative were to turn out to be best analysed as a postposition: as long as the comitative selects an ergative complement, it is the ergative that is triggering the relevant suppletion.

To summarize this section, we have shown that within the realm of case suppletion in personal pronouns, we find the expected patterns of AAA, ABB, ABC. As predicted, we find no ABA patterns within the basic sequences defined by (14), and only sporadic and dubious examples with other oblique cases.²⁷ This supports the view that the grammatical representation of morphological case is based on containment, with oblique cases (including the dative) obligatorily containing dependent and in turn unmarked marked cases. Furthermore, the findings support the proposal that suppletion patterns are sensitive to this internal complexity.

3.6 AAB: suppletion and syncretism

At this point, we turn to the final pattern of case suppletion in our survey, AAB, where suppletion seems to target the ‘third’ (or further) case in a pronoun’s paradigm, rather than suppling immediately from the second case onwards. As noted above, AAB is not attested with degree suppletion. As Bobaljik discusses, however, the basic theory of *ABA (the containment hypothesis, in tandem with the elsewhere condition) does not exclude AAB – Bobaljik (2012) (see also Bobaljik & Sauerland 2018) argued that it is in principle generable, but that additional factors conspire to keep it from arising with degree suppletion. In our survey of case suppletion, we do find cases of AAB.²⁸ Before turning to the theoretical interest, a clarification is in order regarding the two types of patterns that may be described as AAB. Once more, Nakh-Daghestanian languages provide an array of comparative data, this time from the 2SG pronouns, given in Table 20.

Table 20: Syncretic AAB 2SG patterns in Nakh-Daghestanian

	ABS	ERG	DAT	
Aghul	wun	wun	was	{A=A}A
Archi	un	un	wa-s	{A=A}B
Hinuq	me	me	ded-ez	{A=A}B
Tsez	mi	mi	deb-er	{A=A}B

The Aghul, Tsez, Hinuq and Archi pronouns show no distinction between the absolutive and ergative pronouns, but the dative is suppletive relative to these. In our

²⁷Our study encompasses primarily personal pronouns, although other pronoun types (demonstrative, interrogative, etc.) should, all else being equal, show analogous patterns. A potential ABA counterexample comes from Khakas demonstratives (Brown et al. 2003, Baskakov 1975) called to our attention by Stanislao Zompì, though as Zompì notes, it is only problematic if one accepts that there a single, suppletive, demonstrative paradigm, as opposed to two defective series of demonstratives, with overlapping, but slightly different, meaning, cf., perhaps, Baskakov (1975, 151).

²⁸AAB is also found in our number survey, and is frequently attested in suppletion for clusivity, see Moskal (2017).

view, these are, however, not compelling evidence of an AAB pattern. Instead, these are examples of case syncretism: in the 2SG pronouns in these languages, the contrast between absolutive and ergative is simply neutralised. These may be modelled, for example, via impoverishment, deleting the ergative case feature, so that the unmarked case exponent (absolutive) is used in the ergative context. On such a view, these pronouns show simply a two-way contrast (unmarked/direct case vs. dative), which is an AB pattern, and not really a true instance of AAB.

Similar patterns involving syncretism are seen in various Indo-European languages. For example, we see syncretic AAB in German for the third person non-masculine singulars, and in the third person plural (Table 21), whilst this pattern is also seen in Kadugli in Table 22 (Krongo, Reh 1985).

Table 21: {A=A}B in German.

	NOM	ACC	DAT
3SG.M	er	ihn	ihm
3SG.F	sie	sie	ihr
3SG.N	es	es	ihm
3PL	sie	sie	ihnen

Table 22: {A=A}B in Kadugli.

	SUBJ	OBJ	DAT	ALL	LOC
1SG	àʔàŋ	àʔàŋ	àʔàŋ	nkàtí	kàtí
2SG	ùʔùŋ	ùʔùŋ	ùʔùŋ	nkòtú	kòtú
1EXCL	óow	óow	óow	nkòtíg	kòtíg

McFadden (2014, 2018) finds additional support for the importance of treating syncretism separately from suppletion in the realm of non-suppletive stem changes for case in a sample of languages from four families, namely that the only way to produce a (surface) AAB pattern in the languages in his sample is through syncretism of AA cases. The stem alternations McFadden discusses consistently distinguish the unmarked (nominative) from all other cases, with the sole exceptions being instances where the nominative and accusative are syncretic, seen in Latin in Table 23.

If all AAB patterns were this way (i.e. involving full syncretism of the first two cases), then one might explore the possibility that suppletion may be triggered by the mere presence of case, i.e., the first case distinct from the unmarked case suffices to trigger suppletion. However, if we return to Nakh-Daghestanian second person singular pronouns, we see that in Andi, Chamalal, Inxokvari, and Khinalugh, whilst the forms of

Table 23: Case driven stem alternations, (McFadden 2014, 2018).

	NOM	ACC	PART/GEN	INE/DAT	Gloss	
Finnish	ihmi-nen	ihmi-se-n	ihmi-s-tä	ihmi-se-ssä	‘person’	AAA
Latin	it-er	it-er	itiner-is	itiner-i	‘journey’	{A=A}B
Icelandic	mað-ur	mann-Ø	mann-s	mann-i	‘man’	ABB
Tamil	maram	maratt-ai	maratt-Ø	maratt-ukku	‘tree’	ABB

the absolutive and ergative cases are very similar, they are not fully syncretic (Table 24, compare the related languages in Table 20). There is an irregularity in the ergative case that is not present in the absolutive. If the cases were syncretic, then (by definition) there should be no difference between the two of them, as there would be no distinct case feature for the irregular rule to target in order to make them distinct. These are then clearly AAB cases distinct from the {A=A}B cases, and we cannot maintain a view whereby suppletion is always triggered by (at least) the first marked case.²⁹

Table 24: AAB without syncretism in Nakh-Daghestanian 2sg pronouns.

	ABS	ERG	DAT	
Andi	mín	min	du-j	AAB
Chamalal	mì:	mín	du-la	AAB
Inxokvari	mó	me	dub-ul	AAB
Khinalugh	vi	va	oX(ir)	AAB

Another case of AAB without syncretism among the first two cases comes from Wardaman, Table 25 (Merlan 1994). Here, we can see that the difference between the first two cases is not reflected by means of an irregularity in the form, but by the presence of a case suffix in the ergative form that is not present in the absolutive.

Table 25: AAB without syncretism in Wardaman.

	ABS	ERG	DAT
3SG	narnaj	narnaj-(j)i	gunga
3PL	narnaj-bulu	narnaj-bulu-yi	wurrugu

We must therefore conclude that genuine AAB patterns are attested, and ask why pronominal suppletion is different in this way from adjectival suppletion. There are

²⁹Nakh-Daghestanian is a rich source for suppletion. In addition to the A(A)B and AAB patterns discussed, one also finds ABB patterns in among the 2SG pronouns, as in Avar: ABS: *mun*, ERG: *du-la*, DAT: *du-r*.

two possibilities that we see, both relating to how containment is represented. Firstly, case can be a complex feature bundle on a single head, or, secondly, single features that are spread over various nodes in the tree. We are unable to adjudicate between them on the evidence available to us, and so leave them as potential approaches, in the hope that future research will shed further insight on the issue.

The first possibility, pursued in [Smith et al. \(2015\)](#), is that case categories are indeed internally complex, but containment is represented as a complex feature bundle, and not the hierarchical case containment of (14) or [Caha \(2009\)](#). We could therefore assume the case features to be represented as follows (the labels for case features, and the question of whether the unmarked case has any features are not important for the general point):

$$(19) \quad \left[\text{UNMARKED} \right] \quad \left[\begin{array}{c} \text{UNMARKED} \\ \text{DEPENDENT} \end{array} \right] \quad \left[\begin{array}{c} \text{UNMARKED} \\ \text{DEPENDENT} \\ \text{OBLIQUE} \end{array} \right]$$

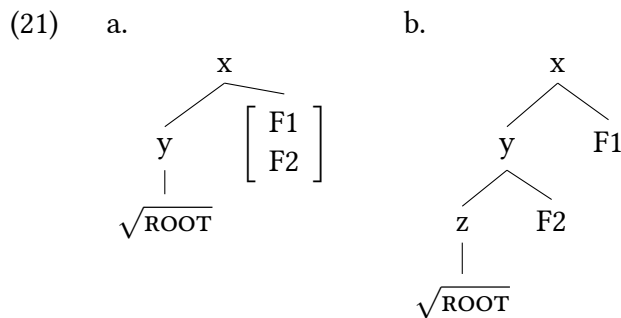
Since each case contains (the features of) all the cases to its left on the hierarchy, the *ABA prediction is maintained. The rules of exponence for the Icelandic 1SG in (15) will have the same effect relative to these representations as they do relative to (14) – the dependent case (accusative) is contained in the obliques (dative) and therefore an ABB pattern, not an ABA pattern, will arise. Similarly, the Wardaman AAB pattern can be readily characterised:³⁰

$$(20) \quad \begin{array}{l} \text{a. } [3, -\text{SG}] \rightarrow \text{wurrugu} / _ \left[\begin{array}{c} \text{ABS} \\ \text{ERG} \\ \text{DAT} \end{array} \right] \\ \text{b. } 3 \rightarrow \text{gunga} / _ \left[\begin{array}{c} \text{ABS} \\ \text{ERG} \\ \text{DAT} \end{array} \right] \\ \text{c. } \left[\begin{array}{c} \text{ABS} \\ \text{ERG} \end{array} \right] \rightarrow \text{-yi/-ji} \\ \text{d. } \text{ABS} \rightarrow \emptyset \\ \text{e. } [-\text{SG}] \rightarrow \text{-bulu} \\ \text{f. } 3 \rightarrow \text{narnaj (elsewhere)} \end{array}$$

Representing complex cases as involving featural, but not structural, complexity would allow the difference between adjectival suppletion (AAB unattested) and case

³⁰(20) represents one possible way of expressing the interaction of number and case in Wardaman, where the non-singular marker *-bulu* is absent in the plural dative. As in any non-transparent containment structure, an additional mechanism is needed to ensure that the ergative exponent *-yi/-ji* is not overtly expressed in the dative. Theories invoking containment have ready means to express this.

suppletion (AAB attested) to be treated as a difference in locality, following the logic set out in (Bobaljik 2012, 158-163). Bobaljik notes that both of the structures in (21) will exclude the *ABA pattern, but that at the same time, if there is a condition of structural adjacency on suppletion (such that the trigger for suppletion must be on the node adjacent to the root), then root allomorphy conditioned by F1 will be possible in (21a) but not in (21b). In the latter, the more marked feature (F1) is not sufficiently local to the root to govern suppletion, since it is not adjacent.



While this may seem to be a simple approach to the difference between case and degree morphology, it relies on the assumption that structural adjacency is a condition on suppletion. There is however emerging evidence, to which we now turn, that such a condition is too strict, and that there are some structures like (21b) in which F1 may condition root allomorphy (Merchant 2015, Moskal & Smith 2016).

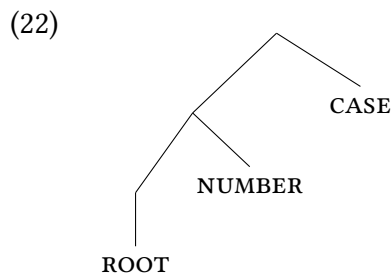
3.7 Adjacency as a restrictor on allomorphy?

Of particular relevance to the present issue are the third person pronominal paradigms in Khakas and Tamil, and interrogative pronouns in Rutul. In Khakas (also discussed in Brown et al. 2003), we see that the pronouns are suppletive in the singular: the base changes from *ol-* in the nominative case to *an-* in other cases. However, in the plural forms, the base is uniformly *ol-*.

Table 26: Blocking of case suppletion in Khakas (Baskakov 1975, 146)

	SG	PL
NOM	<i>ol</i>	<i>olar</i>
ACC	<i>ani</i>	<i>olarni</i>
DAT	<i>aɣaa</i>	<i>olarya</i>
LOC	<i>anda</i>	<i>olar da</i>
LAT	<i>aniɟzar</i>	<i>olarzar</i>
ABL	<i>annaɟ</i>	<i>olardaɟ / olarnaɟ</i>

It seems reasonable to posit that the suppletive root is blocked from arising in the plural forms because the plural suffix *-lar* linearly (and structurally) intervenes between the base and the case suffix, as in (22).³¹ An approach that assumes that allomorphic relations can only be established between elements that are hierarchically adjacent (or linearly adjacent, see Embick 2010) can easily capture this blocking.



However, consider further the first and second person pronouns in Tamil in Table 27.³² Here we see suppletion for case across the plural morpheme *ga(l)*. In the plural form, the dative case morpheme *-ukku* lies outside the number morpheme, and hence is neither linearly nor structurally adjacent to the root. The fact that there is still suppletion in this configuration shows that adjacency cannot be a universal restrictor on allomorphy.³³

Table 27: Suppletion across a number head in Tamil (Asher 1982, 118)

1PERS	SG	PL	2PERS	SG	PL
NOM	naan	naan-ga(l)	NOM	nii	niin-ga(l)
GEN/OBL	en	en-ga(l)	GEN/OBL	on	on-ga(l)
DAT	en-akku	en-gaḷ-ukku	DAT	on-akku	on-gaḷ-ukku

³¹This is consistent with Greenberg’s Universal 39: “Where morphemes of both number and case are present and both follow or both precede the noun base, the expression of number almost always comes between the noun base and the expression of case.”

³²The relevance of these forms was originally pointed out by an anonymous reviewer of Bobaljik (2012). Andrea Calabrese, in work in progress, offers an alternative characterization in which *on-*, respectively, *en-* are the underlying forms of the pronominal bases and in which no suppletion is involved. Rather, the nominative forms involve an augmentation of the base (compare our treatment of Archi, above), mirroring in some ways the historical development of the irregular nominatives for the first person, at least (Andronov 2003, 156-163).

³³Not all cases are shown here. The genitive/oblique is zero-marked, and thus may give the impression that the dative (and other postpositional cases such as the locative, not shown) are built from the genitive/oblique. We take no stand on whether the dative is built from the genitive (since we have remained agnostic about the position of genitive in a case hierarchy) or whether all the obliques abstractly contain the accusative, with a zero marker in the genitive making it look “smaller”. Our discussion here focuses on the relation between case and number.

Another, more minimal, contrast possibly making the same point comes from Chuvash (Turkic, Table 28) and Evenki (Tungusic, Table 29). In Chuvash, we can isolate first person *p*- corresponding to second person *s*-. In the singular, the first person *p*- alternates with *m*- in all the non-nominative cases. But in the plural, where the person formative is followed by the (local person) plural marker *-ir-*, suppletion for case does not obtain and *p*- is retained throughout.

Table 28: Chuvash local pronouns (Clark 1998)

	NOM	ACC/DAT	INSTR	ABL
1SG	epĕ	mana	man-ta	man-ta-n
2SG	esĕ	sana	san-ta	san-ta-n
1PL	epir	pire	pir-te	pir-te-n
2PL	esir	sire	sir-te	sir-te-n

In Evenki, we can similarly isolate first person *b*-, alternating (as in Chuvash) with *m*- in the non-nominatives, and corresponding to second person *s*-. But unlike Chuvash, Evenki shows this suppletion in both the singular, and the plural.³⁴

Table 29: Evenki local pronouns (Nedjalkov 1997)

	NOM	ACC	DAT	ABL
1SG	bi	mine	min-du	min-du-k
2SG	si	sine	sin-du	sin-du-k
1PL	bu	mune	mun-du	mun-du-k
2PL	su	sune	sun-du	sun-du-k

A more convincing minimal pair showing the variation in whether overt number marking is an intervener is provided by Erschler (2017) in his discussion on *wh*-words in Northeast Caucasian languages: in a single language, Rutul (Ibragimov 1978) (as cited in Erschler 2017), ‘what’ displays a blocking effect of an overt (plural) morpheme (Table 30, left), whilst ‘who’ shows suppletion across the same overt (plural) morpheme (Table 30, right).

³⁴Our conclusions from the Chuvash versus Evenki contrast are tentative, not least because (i) the alternation *p/b~m* could be morphophonological, rather than suppletive, and (ii) whether the plural *u* intervenes between the root and the case marker in Evenki depends on how one segments the plural pronominal base. If the pronouns are segmented as *b-i*, *m-i-ne*, *s-u*, *s-i-ne* etc, recognising distinct person and number morphemes, then the *b~m*- alternation has a non-adjacent trigger (case). Alternatively, one could posit an ablaut rule, changing *i* to *u* without decomposing the pronominal bases into person and number, which would leave the case-driven alternation as applying to structurally adjacent morphemes.

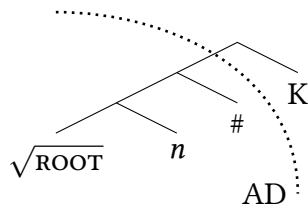
Table 30: Suppletion in Rutul wh-words (Erschler 2017).

‘what’			‘who’		
	SG	PL		SG	PL
NOM	šiv	šiv-dəbər	NOM	vuš	vušer
ERG	hid-iræ	šiv-dəbiš-æ	ERG	hal-a	hal-dəbiš-æ
GEN	hid-id	šiv-dəbiš-də	GEN	hal-də	hal-dəbiš-də

It appears, then, that in order to explain the differing distribution of AAB patterns in adjectives (unattested) versus pronouns (attested), locality in morphology, like in syntax, may need to appeal to interveners and thus, perhaps domains, rather than (structural or linear) adjacency. In principle, a node which occurs between the target and trigger of suppletion (F2 in (21b)) may or may not be an intervener, blocking some interaction between F1 and the root. Evidently, the comparative node in (1) intervenes to block suppletion of the adjectival root triggered by the superlative (thereby excluding AAB patterns), but number in (22) is not always an intervener. There are various theories of locality and interveners in the current literature. In some theories, the intervening nodes constitute domain boundaries, perhaps analogous to phases in syntax. Although we do not aim to adjudicate among competing theories on this point (see Moskal 2015a for a review), we offer a few relevant observations about how the data here bears on competing approaches. In the worst case, whether the context for suppletion requires an adjacent or merely c-commanding trigger could be stated in the contexts for the individual rules of exponence introducing suppletive allomorphs. The item-specific variation in Rutul suggests such an approach may be needed.

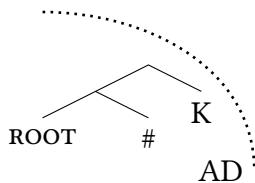
One approach which may draw the right cut between adjectives (no AAB for comparison) and pronouns (AAB for case), at least to a first approximation, is that developed in Moskal (2015a,b). Moskal argues for a dynamic definition of cyclic heads (see Bobaljik & Wurmbbrand 2005), and parasitic on this, a definition of an *accessibility domain* (AD): following Marantz (2000, 2007) and others, she proposes that the functional heads that categorise roots are potentially cyclic, but define a cyclic domain only if they constitute the highest in a particular sequence of projections. The accessibility domain for a root consists of the heads that have been merged into the derivation when the cycle containing that root is fixed – thus, the AD for a root will contain the first category-defining node above the root, and one node above that (since that node determines that the potentially cyclic category-defining node is in fact the cyclic node). This is illustrated for a noun in (23).

(23)



Moskal argues that for lexical nouns, this has the effect that case information is too far away from the root in order to factor into allomorphy. Pronouns on the other hand are deficient; Moskal argues that they lack category defining nodes, and so there is no domain created low in the structure that contains just the pronominal base. Thus, case information is accessible to the root, and suppletion for case is possible in pronouns.

(24)



This same approach, which defines a larger accessibility domain in pronouns than in nouns, may be brought to bear to exclude AAB patterns in adjectives, if adjectives, like nouns, have a category defining node (*a*) between the root and the comparative affix. Just as *K* will be too far away from the root to trigger root suppletion in (23), so will the superlative head be too far away to trigger root suppletion in adjectives, while AAB patterns in pronouns will be permitted, as there is no internal domain in pronominals.³⁵

³⁵The apparent ‘blocking’ effect seen in Khakas is not a locality effect under this approach and must be stated in the vocabulary insertion rules of that language. Moskal & Smith (2016) propose that it is the non-nominative singular forms that are suppletive, and are picked out by VI-rule in (ia) that makes reference to both number and case. All other forms (nominative singular and all plural forms) use the elsewhere form of the base, determined by the elsewhere rule in (ib):

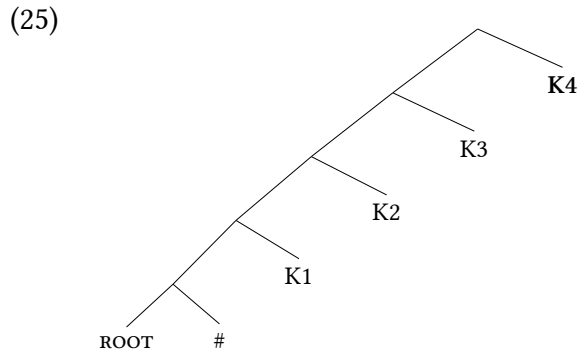
- i a. [3] ⇔ an / _] SG] K]
- b. [3] ⇔ ol

Alternatively, one may simply state in the rule itself that the Khakas non-nominative form requires adjacency to *K* (as in (iia)) as opposed to the Tamil oblique allomorph, which requires only (domain-local) *c*-command, but not adjacency (iib). If singular number is pruned or otherwise not present in the structure at the point of vocabulary insertion, the rules in (ii) will distinguish the two types of system.

- ii a. [3] ⇔ an / _] K]
- b. [2] ⇔ on / _] ...] K]

Since the blocking effects are not immediately relevant to our purposes, we refer the reader to Moskal & Smith (2016) for further discussion.

Under Moskal’s approach, positing an articulated structure for case (multiple K heads) would not change the fundamental asymmetry between nouns and pronouns. All the K heads would be outside the domain of a lexical noun in (23), and unable to condition root suppletion, while in pronouns, since there is no domain delimiter in (25), all case information will be accessible to the pronominal base.



We may then maintain the structural containment view of case, roughly in line with [Caha \(2009\)](#) (although as noted above, without a commitment to a unique total order among the obliques). This also obviates problems when there is overt containment of cases, as well as keeping a strict parallel between case and adjectival suppletion.³⁶ If adjacency is not a universal condition on suppletion, then it becomes possible to maintain the structural representation of containment in (14). Case and degree may have analogous structural representations, but the difference would then have to lie in whether or not there is a locality domain.

In sum, the fact that genuine AAB patterns are found for case is not surprising, given the different options for capturing them that we have outlined in this section, particularly the relative lack of structure found in functional items like pronouns when compared to lexical items like adjectives. One possibility is that containment of case can be represented as involving complex features on a single node, however this means giving up on the strict parallel between case and degree suppletion. Such a strict parallel can be maintained by representing case containment over distinct K projections, coupled with an independently necessary relaxation of adjacency requirements in allomorphic relations. Crucially however, the logic of containment coupled with the Elsewhere Condition continues to rule out ABA patterns.

³⁶ It should be noted that adopting this view of case containment may yet turn out to be inconsistent with the view of locality advocated for in [Moskal \(2015a\)](#). There, she argues that a small number of instances of case suppletion in lexical nouns results from the absence of a number node, which brings case into the Accessibility Domain of the root. However, adopting the structural containment of case means that in the ‘one-node-above-cyclic-nodes’ approach that Moskal gives, case suppletion in lexical nouns is unable to be stated, since the only node able to be targeted would be K1, and hence there would be no way to distinguish K1 from K2. A similar set of questions is raised if NumberP is split, as we suggest below, or if there are other functional elements in the nominal spine.

3.8 Summary

In this section we have seen evidence for two major claims. Firstly, the patterns from suppletion lend support to the proposal made in various places that morphological case is complex, with more complex cases containing their less complex counterparts. Thus, at least to a first approximation, our findings are mostly in accordance with the proposals like those of [Caha \(2009\)](#) in which the (or a) case hierarchy is formally represented in the grammar such that oblique cases are built on the dependent case, which in turn contains the unmarked case. However, we have left open the precise nature of that representation, in particular, whether it involves structural or featural containment. The second finding of our study is that it appears as though we can generalise the model in [Bobaljik \(2012\)](#) for capturing possible suppletion patterns to an independent empirical domain.

4 Number driven suppletion in personal pronouns

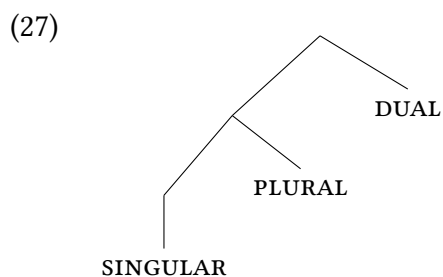
With the facts from case in mind, we now turn our attention to number. Given that we are looking to see whether suppletion can ever revert to a default form as the category becomes more complex, simply looking at the familiar cases of singular-plural number suppletion like *person* ~ *people* will not suffice. Therefore, we must look at languages which have at least a three-way number distinction, for instance languages which have a dual in addition to singular and plural.

4.1 Complex number

As with case, we take statements of typological markedness as our starting point. For number, the relevant observation is the following:

- (26) No language has a dual unless it has a plural. (from Universal 34, [Greenberg 1963](#), [Corbett 2000](#))

Postponing for the moment a more refined understanding of the categories involved, we might assume, as we did for case, that the markedness hierarchy is reflected as structural containment, as in (27):



Similarly, when looking for direct evidence of containment in the overt morphology, we find examples that are straightforwardly consistent with (27), i.e., in which the form for the dual transparently contains the form for the plural.³⁷ This is seen in the Manam demonstrative (Lichtenberk 1983), where we can see the plural morpheme *-di* is also contained in the dual form, which is formed from the plural form with the addition of a linker morpheme and the dual suffix *-ru*.

- (28) a. áine ŋára b. áine ŋára-di c. áine ŋara-dí-a-ru
 woman that-SG woman that-PL woman that-PL-LINKER-DL
 ‘that woman’ ‘those women’ ‘those two women’

Furthermore, looking at systems that are three-way contrastive for number, other than for SINGULAR – PLURAL – DUAL, we again find containment patterns. For instance, Warrwa has a MINIMAL – AUGMENTED – UNIT AUGMENTED system (McGregor 1994, 20), and here we see that the unit augmented form is built on the (suppletive) augmented form, as in Table 31.³⁸ This system differs from the SINGULAR – PLURAL – DUAL system in that minimum means the ‘the logical minimum’, augmented ‘more than the logical minimum’ and unit augmented ‘the logical minimum plus one’; see our discussion of Ilocano below, and Corbett (2000, 166-169).

Table 31: Pronominals in Warrwa

	MINIMAL	AUGMENTED	UNIT-AUGMENTED
1EXCL	ngayu	yaarra, yarrin	yaarra-wili, yarrin-bili
1INCL	yawu	yaddirr, yarru	
2	juwa	kurra	kurra-wili, kurrawa-wili
3	kinya	yirra	yirra-wili

As before, the structure in (27) leads us to expect that ABB patterns should be possible, but ABA should be unattested. Whether ABC and AAB patterns should exist depends on the additional questions of locality, and whether number containment is represented as structural containment, as in (27), or featural containment, as in (21), topics we return to as we refine the discussion below.

4.2 Number suppletion

For pronominal number suppletion we looked at an initial sample of 80 languages, which was supplemented with information from the extensive Free Personal Pronoun

³⁷In fact the opposite is also attested, with the plural apparently containing the dual. For expository reasons, we hold that in abeyance for the moment, returning to such evidence in section 4.3.

³⁸According to McGregor, the inclusive does not have a specific unit-augmented form.

System database compiled by Norval Smith (Smith 2011). We also utilised the suppletion database at the Surrey Morphology Group (Brown et al. 2003). Table 32 summarizes our results, where in the triplet XYZ, X = singular/minimal, Y = plural/augmented, and Z = dual/unit-augmented.

Table 32: Number suppletion: Summary

Pattern	Prediction	<i>n</i> attested	Representative languages
AAA	✓	numerous	Mapuche, Dumi
ABB	✓	48	Kayardild, Kham
ABC	✓	19	Savosavo, Pitta-Pitta
AAB	?	3	Wambaya, Yagua
ABA	✗	(1)	Yagua?

The first thing of note about the attested patterns is that all the attested patterns, and crucially also the unattested patterns, are in accordance with Bobaljik’s findings regarding degree suppletion. That is, we find examples where the base remains constant (AAA), cases where the base suppletes once, but the dual and plural share a common base (ABB), and further cases where the base suppletes twice and the singular, plural and dual all have different bases (ABC).³⁹ We also find examples of ABB patterns. What we do not find in our sample are any clear instances of ABA suppletion. Like the Archi 2PL in the case section, there are some examples of ABA-like patterns worthy of deeper discussion, most notably the Yagua third person, but as we discuss below, we find them to be far from clear-cut.

As with the case dataset, we find some paradigms that are hard to classify, but for which ABA is thankfully not among the plausible options. An example is given in Awtuw, in Table 33.

Table 33: Awtuw pronouns (Smith 2011)

	SG	PL	DL	
1	wan	nom	nan	ABB
2	yen	om	an	¬ABA
3M	rey	rom	ræw	AAA
3F	tey	rom =3MPL	ræw =3MDU	SYN:ABB

The Awtuw second person looks on its own like a candidate for an ABC pattern, but looking at the system as a whole, and the first person in particular, it is clear that *om* is

³⁹See Daniel (2005) for an overview of plural marking in independent pronouns. In Daniel’s survey of 261 languages, almost 3/4 show suppletion for number, either with (69) or without (114) an independent plural affix. Daniel does not include duals, and so is not informative for the current study.

the (historically expected, [Palmer 2016](#), 272-3) plural marker, while both *om* and *an* are shared with the first person. One could treat *om* and *an* as just number markers, positing a second person non-singular \emptyset base corresponding to first person *n*- (compare the discussion of Polish clitics above), and thus an ABB pattern. We have remained agnostic, indicating Awtuw second person as \neg ABA (“Not ABA”) in our dataset (online).

The Awtuw third persons illustrate another common occurrence: there is a gender contrast in the singular only: all non-singulars use the masculine gender marker *r*-. While the masculine dual and plural are thereby used as the non-singulars corresponding to the feminine singular, we do not treat these as suppletion for the feminine. We indicate these in the online appendix, as here, as “SYN:ABB”, that is, an apparent ABB pattern, but one that arises via syncretism (neutralisation) rather than as suppletion. Such syncretic patterns are not counted in our pattern tallies.

Another set of issues that is specific to number suppletion is how to deal with clusivity. Two points are worthy of note. The first is illustrated by Djamindjung in [Table 34](#). Djamindjung, like many languages, draws an inclusive/exclusive distinction in the first person, but only in the non-singular numbers. For languages with such patterns, we have taken the 1SG to count as the corresponding singular for both the inclusive and exclusive series. For Djamindjung, both the inclusive and exclusive are suppletive relative to the singular, yielding an ABB and an ABC pattern, respectively. In other languages, one or both of the non-singular series may share a root formative with the first person singular. For a more detailed study of suppletive patterns in clusivity contrasts, see [Moskal \(2014, 2017\)](#).

Table 34: Clusivity in Djamindjung ([Smith 2011](#))

	SG	PL	DL	
1EXCL	ɲayug	yirri	yirrinji	ABB
1INCL		yurri	mindj	ABC
2	nami	gurri	gurrinji	ABB
3	dji	burri	burrinji	ABB

The second remark concerns minimal-augmented number systems. The distinction is primarily relevant only for the first person inclusive pronouns (see [Corbett 2000](#), 166-169). While a singular inclusive is a contradiction, a minimal inclusive, like Warrwa *yawu* in [Table 31](#), denotes the (unique) speaker and one (unique) hearer, the logical minimum number of referents that can include both speaker and hearer. Such pronouns are often described as “first person inclusive dual” pronouns. For languages that have no other instantiations of a dual number, this gives the appearance of a three-way number contrast in the first person only. On the other hand, a minimal-augmented analysis of the same facts treats such a system as having only a two-way

number contrast throughout the language. The two analyses of Ilocano, from an early discussion of these issues (Thomas 1955), are given in Tables 35 and 36.

Table 35: Ilocano as a language with dual

	SG	DL	PL
1EXCL	co	ta	mi
1INCL			tayo
2	mo	–	yo
3	na	–	da

Table 36: Ilocano with a Minimal-Augmented analysis

	MIN	AUG
1EXCL	co	mi
1INCL	ta	tayo
2	mo	yo
3	na	da

On the analysis in Table 35, Ilocano would end up having an ABB pattern, if *ta* is the dual of first person *co*, but on the analysis in Table 36, there is only a two way number contrast, and the language is not relevant to our central prediction. We have treated all languages in our sample which, like Ilocano, have a dual number uniquely in the first person as amenable to a Minimal-Augmented analysis, and excluded them from our counts.

The absence of ABA patterns points, as above, to the conclusion that the distribution of suppletion is not random, but follows consistent patterns across languages and domains. The absence of ABA patterns in particular reinforces our contention that features are hierarchically structured, even within small functional categories like pronouns. Before turning to some complications, we present a quick overview of our empirical findings.

4.2.1 AAA, ABB, etc

There is, of course, no requirement that pronouns show suppletion for number, thus AAA patterns are well attested. Mapuche (Smeets 2008) and Northern Qiang (LaPolla & Huang 2003, 50-54) serve to illustrate:⁴⁰

Table 37: AAA in Mapuche (Smeets 2008)

	SG	PL	DL
1	iñché	iñchiñ	iñchiu
2	eymi	eymün	eymu
3	fey	feyengün	feyengu

Table 38: AAB and AAA in (Northern) Qiang (Smith 2011)

	SG	PL	DL
1	qa	təile	təizzi
2	ʔü	ʔile	ʔizzi
3	the:	themle	thizzi

⁴⁰We also note that Mapuche builds the plural from the dual, not vice versa.

In Qiang, we take the first person to show an ABB pattern, assuming that no phonological process relates *qa* to *tci*, but we see consistent formatives across number, thus AAA patterns, for the second (*?ũ/i-*) and third (*the/i-*) persons, modulo the vowel alternations.⁴¹

Where suppletion is attested, by far the most common pattern is ABB, where the plural form and the dual form share a suppletive base that is different from the base of the singular. We illustrate with examples from Kayardild (Evans 1995), and Kham (Watters 2002), in Table 39 and Table 40, respectively. In both Kham and Kayardild, we can see that both the dual and the plural forms are suppletive with respect to the corresponding singular forms of the pronouns. In Kayardild, both appear to decompose straightforwardly into a non-singular person formative: second person *ki-*, third person *bi-*, and an element that reflects number. ABB is also seen in Gothic second person pronouns (Smith 2011), which we illustrate in Table 41.

Table 39: ABB in Kayardild (Evans 1995)

	SG	PL	DL
2	nyinka	kilda	kirra
3	niya	bilda	birra

Table 40: ABB in Kham (Watters 2002, 160)

	SG	PL	DL
1	ŋa:	ge:	gin
2	n:	je:	jin

Table 41: ABB in Gothic (Smith 2011)

	SG	PL	DL
2	þu, þuk	jus, izwis	jut, igqis

Turning to ABC patterns, there are various candidates found in our sample. Firstly, in Kham, while the personal pronouns provide ABB patterns as just noted, the reflexive pronouns (and possessive prefixes) constitute ABC patterns as in Table 42.⁴² Next, in Jehai (Austro-Asiatic) second and third person pronouns show an ABC pattern:⁴³

ABC patterns are seen in the second and third masculine pronouns in Savosavo (Table 44) and in Bukiyip second person (Table 45).

Finally, AAB patterns were attested in our survey, though they were rarer than the other patterns. Overall, we identified three candidates for AAB patterns, which are listed in Table 46.⁴⁴

⁴¹Third person *the* is also a demonstrative, but has a different plural and dual as demonstrative than as pronoun.

⁴²Thanks to Kenyon Branan for pointing these out to us.

⁴³The neutralization of a 2 vs. 3 person contrast in the plural suggests that only one of these is

Table 42: ABC in Kham reflexive pronouns (Watters 2002, 162-3)

	SG	PL	DL
3 REFL	ol	ya:	ni:
3 POSS	o-/u-	ya-	ni-

Table 43: ABC in Jehai (Smith 2011)

	SG	PL	DL
2	miʔ, mɔʔ, paj	gin	jih
3	ʔoʔ	gin	wih

Table 44: Savosavo (Smith 2011)

	SG	PL	DL
2	no	me	pe
3M	lo	ze(po)	to

Table 45: Bukiyip (Smith 2011)

	SG	PL	DL
2	nyak	ipak	bwiepú

As in the discussion of case, the existence of AAB patterns bears on the question of locality for suppletion with respect to the representation of containment. We will revise our assumptions about (27) below and return to AAB patterns at that time.

In our survey of patterns of pronominal suppletion for number, AAA, ABB, ABC and AAB were the only patterns we found. We did not find any clear ABA patterns, confirming the basic prediction of a structural approach to suppletion. For the sake of completeness, we note that in our survey there were a few examples that could be classified as ABA at first blush, though we suggest that these are best analysed in other ways. We turn to these now.

In Biri (Table 47, Pama-Nyungan Terrill 1998, Smith 2011) there is a hint of an ABA pattern in the second person if the relevant formative is *yi-*, *yu-*, *yi-*.

But there are two alternatives. On the one hand, since the paradigm as a whole is irregular, one could treat each form as memorised whole, thus as ABC patterns for both first and second person.⁴⁵ But this seems to us to miss the obvious generalisation that initial *y-* (or perhaps *yu-* with a rule changing the vowel) is common to all and only the second person pronouns, just as *d-* is in German.⁴⁶ Likewise, all and only the first person forms begin with *ŋa*. Recognizing a consistent formative for both the first

properly considered an ABC pattern.

⁴⁴Sources: Wambaya (Nordlinger 1998), Yagua, (Payne & Payne 1990), Dehu (Smith 2011). Smith draws on an old source, and may give an incomplete paradigm. The description in Lenormand (1999, 24-27) decomposes the pronouns into an honorific prefix, a person root, and a number suffix, and presents a more regular picture, with ABB in the first person, but regular AAA person formatives in the second and third persons.

⁴⁵There is no dual in third person.

⁴⁶See Terrill (1998, 23-25) for further discussion of the Biri forms. Terrill suggests an etymology for dual *yibala* that involves “the /u/ being fronted to /i/ after the /y/” (p.25). She suggests also that the alternative second person plural *yubala* is a recent addition to the language. Note that *-bala* is not a regular number affix in the language.

Table 46: AAB patterns for number

		SG	PL	DL
Wambaya	1INCL	ngawu(rniji)	ngurruwani	mrindiyani
Yagua	2	jiy	jiryéy	sááda
Dehu	3M	angeice	angate	nyido

Table 47: The pronominals of Biri (Terrill 1998, Smith 2011)

	SG	PL	DL
1	ɲaya	ɲana	ɲali
2	yinda	yura/yubala	yibala
3	nhula	dhana	–

and second persons, we therefore treat Biri as having AAA patterns.

A second tricky data point, and potential candidate for an ABA pattern comes from Yagua (Payne & Payne 1990), with the paradigm given in Table 48. The third person is of interest this time. In the Yagua third person, we find an initial *n-* in the singular and dual, missing in the plural. Comparison of the second and third person duals suggest that the *n-* in the dual is indeed a third person formative. But Yagua differs from Biri in a potentially important way. In Biri, it was possible to identify a formative that occurs in all and only the second person forms, which we take to be strong evidence which the child is attuned to. In Yagua, however, the *n-* meets neither criterion: it does not occur in all the third person forms, nor is an initial *n-* limited to third person forms, since it occurs also in 1EXCL.PL and 1EXCL.DL. We note that it is possible that it is, ultimately, a challenge for our analysis, but it is by no means a robust counter-example.

Table 48: Potential ABA in Yagua (Payne & Payne 1990)

	SG	PL	DL
1EXCL	raay	núúy	nááy
1INCL	ray	ruuy	ruuy
2	jiy	jiryéy	sááda
3	níí	riy	naada

The final potentially challenging pattern comes from two related languages, Nyamal and Wajarri, and so we treat this as one data point, presuming the forms are cognate. We illustrate with Wajarri. The Wajarri pronoun data are given in Table 49, with the pronouns of interest coming from the third person.⁴⁷

It seems as though the third person plural form is suppletive relative to the dual and singular forms. However, it is not clear that it is accurate to consider the third person

⁴⁷We will assume that the explanation give for Wajarri is the same for Nyamal.

Table 49: Wajarri (Smith 2011)

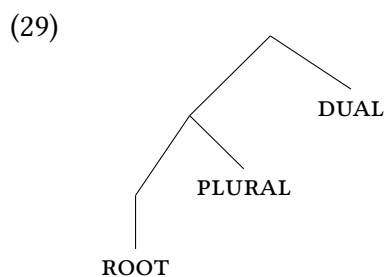
	SG	PL	DL
1EXCL	ngatja	nganju	nganju
1INCL	–	ngantju	ngalltja
2	njinta	njurra	njupali
3	palu(-tja)	tjana	pula(-tja)

forms as constituting a single pronominal paradigm. Dixon (1989, 357) notes that “it is not uncommon for Australian languages to have forms that can be recognised as 3DL and 3PL pronouns, but nothing that could properly be called 3SG”. Instead, this function is often taken up by demonstratives. In Wajarri, it seems plausible to treat the 3rd person non-plural forms as being a part of the demonstrative paradigm. Douglas (1981, 223) notes that the non-plural forms, but not the plural, show a demonstrative-like alternation: the forms *palu* and *pula* “refer to 3rd person singular and dual (respectively) within the local group. To refer to a third person (singular) outside the group *palutja* is used. To refer to third person dual outside the group *pulatja* is used.” In addition, the [pVIV] series includes the positional pronoun *pala*, for which Douglas (1981) gives the gloss ‘that mid-distant person or thing.’ That the non-plural forms are subject to changes in location of the referent, but the plural form is not, suggests that the non-plural forms are part of the demonstrative paradigm and that it is inaccurate to represent all three numbers as (suppletive) forms of a single pronoun.

In sum, while acknowledging the patterns just described, we maintain that it is nevertheless true, for number, that we find extremely clear-cut examples of ABB, ABC and AAB patterns, alongside AAA. We do not find any unambiguously robust evidence of ABA patterns.

4.3 Number: beyond the basics

The data from pronouns as presented are clearly consistent with an analysis where the representation of the dual contains that of the plural, as we discussed above in (27), repeated below:



At this point, three related issues arise. First, our labels PLURAL and DUAL do not do justice to the significant literature on the semantics of number, and our proposed structure should receive further scrutiny from that perspective. Second, although we provided the Manam example in (28) as transparent support of (27), we noted there that the opposite pattern is also attested (footnote 37). Third, and most pressingly, in the course of our investigation of pronouns, we found sporadic examples of suppletion for number in lexical nouns which appear to show the opposite pattern with plural the odd one out, and dual and singular sharing a stem.⁴⁸ The four examples we have found are given below. According to (27), this would constitute an ABA pattern.⁴⁹

Table 50: Number suppletion in lexical nouns

Language	SR	DL	PL	Gloss
Hopi	wùuti	wùutit	momoyam	‘woman’
Lavukaleve	vo’vou	vo’voul	tulav	‘boy’
Slovenian	člòvek	člòvek-a	ljudj-e	‘person’
Yimas	panmal	panmalc-rm	pay-um	‘man’

We suggest that the three issues just raised are interrelated, and that a more sophisticated representation of number than that in (27) allows sufficient flexibility to describe the attested patterns, while still providing the means to exclude the unattested patterns under the general containment logic that unifies the various domains we have investigated. We approach the argument in steps, noting that there are various ways to cash out the ideas presented here, and we present only one as a demonstration that it is possible to do so.

4.3.1 Representing number

We start our reevaluation of (27) with the observation put aside above (footnote 37) that, in the realm of overt containment relations, sometimes the dual contains the plural, but sometimes the plural contains the dual (Corbett 2000, Harbour 2014). Harbour (2014) for example provides striking minimal contrasts from related languages (some of which have more than a three-way number distinction). Sursurunga and Mokilese emphatic pronouns both show a range of number contrasts, but they differ regarding which of the non-singulars has no overt exponent; in Sursurunga (Table 51), the plural serves as the base for the other non-singulars (like Manam), but in Mokilese, it is the

⁴⁸Suppletion for number also occurs with verbs (Veselinova 2006, Bobaljik & Harley 2017) and adjectives (Harbour 2008 on Kiowa), which are beyond the scope of our inquiries.

⁴⁹The reason for why the order of the columns has been switched to SINGULAR – DUAL – PLURAL will become apparent shortly.

dual that serves this function (Table 52).⁵⁰ The pattern of apparently forming plurals from duals is also found outside of Austronesian, for example in the Pama-Nyungan language Panytyima, where the plural form appears to come from the dual, with the addition of *-kuru* (Table 53).

Table 51: Sursurunga: plural in dual (Harbour 2014)

	SG	PL	DL	PCL
1EXCL	ia <u>u</u>	gi <u>m</u>	gi-ur	gi-m-tul
1INCL		gi <u>t</u>	gi-ar	gi-tul
2	iá <u>u</u>	ga <u>m</u>	ga-ur	ga-m-tul
3	-i/on/ái	di	di-ar	di-tul

Table 52: Mokilese: dual in plural (Harbour 2014)

	SG	DL	PL
1EXCL	ngoah	kama	kama-i
1INCL		kisa	kisa-i
2	koah	kamwa	kamwa-i
3	ih	ara/ira	ara-i/ira-i

Table 53: Panytyima: dual in plural (Smith 2011)

	SG	DL	PL
1EXCL	ŋatha	ŋaliya	ŋaliyakuru
1INCL	ŋatha	ŋali	ŋalikuru
2	njinta	nhupalu	nhupalukuru

This observation is relevant, since our theoretical apparatus provides only contingent predictions: the containment structure provides the order of columns, relative to which we predict *ABA. If the containment pattern is SINGULAR – PLURAL – DUAL, as we have been assuming, then the plural should not supplete alone, but if SINGULAR – DUAL – PLURAL is a possible containment pattern, as affix order in Mokilese and Panytyima seem to suggest, then *ABA relative to that structure would exclude suppletion of the dual alone; suppletion of the plural alone would be, relative to that structure, an AAB pattern, and hence admissible.

We will revise this presently, but holding the broader implications of flexible ordering in abeyance for the moment, the problematic patterns in Table 50 would be correspondingly unproblematic if these languages have the Mokilese-like containment pattern. The available evidence is consistent with this view. Hopi, and perhaps Lavukaleve, have nominal affix orders in which the plural appears to be derived from the dual (and Slovenian and Yimas do not provide contradictory evidence).

In some Hopi nominals, the dual and plural are formed by suffixes, *-t* and *-m*. One class of nominals (including some deadjectival forms) mark the dual with one of these

⁵⁰Examples are presented with Harbour’s segmentation and analysis.

suffixes, and the plural with the dual form plus reduplication. In these nominals, including the forms for ‘donkey’ and ‘child’ in Table 54, the form of the plural apparently overtly contains the form with the dual suffix, the reverse of Manam.⁵¹

Table 54: Dual containment in Hopi

SG	DL	PL	
sino	sino-t	sino-m	‘person’
mooro	mooro-t	moo-moro-t	‘donkey’
tsay	tsaayo-m	tsaa-tsayo-m	‘child’

We see this first step in our reevaluation of number as promising. When our survey is expanded to include nouns, as well as pronouns, we find variation in whether the dual or plural is singled out in suppletive patterns. This would be problematic if all languages had the unique structural containment pattern in (27), as implied by Greenberg’s Universal 34. On the other hand, overt containment patterns among plural and dual are known to vary independently. Since our theory of suppletive pattern makes only contingent predictions, relative to a given structure, the observed variation in structure is expected to correlate with the observed variation in suppletion, and the evidence is consistent with this.

At the same time, this interim conclusion is unsatisfactory in two ways. First, allowing for variation in (27) seems to be at odds with Greenberg’s Universal 34, which does not show corresponding variation. Many languages make only a plural versus non-plural distinction, but no known language marks only a dual versus non-dual distinction, which would perhaps be expected if containment structures could vary. In addition, our naive approach to number, using the terms dual, plural, etc. does not match up as well with established theories of number in natural language. Work that looks into the representation of number, and how it relates to *plural*, *dual* etc., such as

⁵¹In Lavukaleve, it appears that there is language-internal variation on this point. In pronouns, descriptively, dual forms are built from plurals. Nouns generally do not show overt containment, however there are some irregular nouns that in the plural end in *lav* (our example above is one of these). It is possible here potentially to decompose the ‘plural’ suffix into *l+av*. *l* is a frequent dual marker in the language, and *av* is a marker of plurality as well, in some words, which is a variant of a general [Vv] morpheme for plurality. In this instance, in terms of the analysis of number to be adopted below, it is possible to view *l* as the spell-out of [–singular], and *av* as the spell-out of [+augmented]. *Tulav*, our example listed in Table 50, would then have the decomposition as follows:

- i. tu -l -av
 √BOY [-singular] [-augmented]

On this analysis, the plural is built on top of the dual for nouns of this type (at least: the only overt morphological evidence we have is for a noun vs. pronoun contrast), and the triple would constitute an AAB pattern, rather than ABA.

Noyer (1992), Harbour (2008, 2011) has converged on the idea that number is complex: not made up of privative features that correspond to plural or dual but rather composed of the features $[\pm\text{singular}]$ and $[\pm\text{augmented}]$. Harbour (2014) in particular shows that a feature system that is based on $[\pm\text{singular}]$ and $[\pm\text{augmented}]$ generates only the attested values of number found across natural languages, whereas an approach such as the one above would overgenerate, being essentially open ended, allowing for distinctions above trial, which are not attested.⁵² The features $[\pm\text{singular}]$ and $[\pm\text{augmented}]$ are semantically defined as in (30) (when a feature has a minus value, the definition is negated).

- (30) a. $[\text{+singular}] = \lambda x[\text{atom}(x)]$
 b. $[\text{+augmented}] = \lambda P . \lambda x: P(x) . \exists y [P(y) \wedge y \sqsubset x]$

The semantics of number is not our focus here, so for a more in depth discussion of these features and how they relate to the wider typology of number, we refer the reader to the cited works and references therein. It suffices to note that $[\text{+singular}]$ has its intuitive value of a quantity of Xs for which no subpart is an X (true of singulars, but not true of plurals), whilst $[\text{+augmented}]$ is true only when the quantity is more than the minimum needed to satisfy the denotation of what the feature applies to.⁵³ What is important here is how these features combine to produce the number values of singular, dual and plural. All three are formed by a combination of these features:⁵⁴

- (31) a. $\text{singular} = [\text{+singular}, -\text{augmented}]$
 b. $\text{dual} = [-\text{singular}, -\text{augmented}]$
 c. $\text{plural} = [-\text{singular}, \text{+augmented}]$

We suggest that an adaptation of Harbour’s approach to number allows us to capture, at the same time, the invariance of Greenberg’s Universal 34, and the variation in affix order and corresponding patterns of suppletion. The first step in this reasoning is to represent the Noyer-Harbour features as instantiating a containment relation, thus we state the hypothesis in (32).

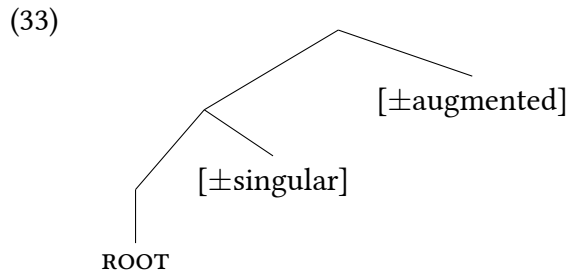
- (32) Number containment hypothesis:
 $[\pm\text{augmented}]$ always contains $[\pm\text{singular}]$.

⁵²This is a simplification of Harbour’s conclusions, which are broader than applying only to languages which make a distinction between singular, plural and dual.

⁵³An alternative to $[\pm\text{augmented}]$ is its inverse: $[\pm\text{minimal}]$. Harbour (2014) settles on $[\pm\text{augmented}]$ since recursion of this feature allows him to capture richer number distinctions including paucals, see Harbour (2014) for other number systems. Depending on the combination of the features, we make further predictions about suppletive patterns where the features stand in containment relations.

⁵⁴Note that the fourth combination $[\text{+singular}, \text{+augmented}]$ is semantically incoherent; $[-\text{augmented}]$ is therefore redundant in the context of $[\text{+singular}]$.

There are a number of ways that (32) could be implemented. For example, like with degree morphology, we could assume that (32) is structural in nature. That is, the functional head Num is in fact more articulated than is usually assumed, and that each of the features [\pm singular] and [\pm augmented] constitutes a head in its own right, as in (33):



Note that because of the way that Harbour sets up the semantics of the features, [\pm singular] must compose with the pronominal root before [\pm augmented] does (Harbour 2014, 206) – composing [\pm singular] after [\pm augmented] would be either vacuous or uninterpretable.⁵⁵ More so than with case, it is therefore not implausible to think of (33) as a consequence of the Complexity Condition (Bobaljik 2012, 212) that, by hypothesis, motivates the containment structure in degree morphology. That is, it could well be the case that learners are forced to posit a structure like (33) for languages with a singular-plural(-dual) contrast, as otherwise there would be too much information in a single head.⁵⁶

The revised structure in (33) has many of the properties that our naive structure had. Among other properties, it faithfully encodes the content of Greenberg’s Universal 34 (26): the contrast between dual and plural ([\pm augmented]) is a subdivision of the non-singulars, thus, a language must first divide the space into singular vs. non-singular in order to make further subdivisions. But does (33) satisfy the containment hypothesis? The answer is a qualified yes, where the qualifications provide just enough flexibility to address the issues raised at the top of this section.

⁵⁵Harbour’s 2008 analysis of adjectival suppletion in Kiowa seemingly requires that [\pm singular] and [\pm augmented] be on the same head, Number⁰. However, that argument relies on the assumption that the trigger for suppletion must be strictly adjacent (structurally and linearly) to the target, an assumption that we have argued above is unsupportable. Note that having both features on a single head also requires a less transparent mapping from syntax to affix order, when both [\pm singular] and [\pm augmented] have discrete exponents, as in Manam. For whatever it is worth, our proposal will allow a 1:1 mapping from syntactic heads to overt affixes, respecting some version of the mirror principle (Baker 1985). In more recent work, Harbour does distribute the features across nodes, for example, in the analysis of constructed duals in Harbour (2017).

⁵⁶It should be borne in mind that we are not making the claim that this is the universal structure of NumP. Harbour (2014) shows that there are languages that do not make use of the feature [\pm singular], and only use [\pm augmented] (languages which only make a minimal-augmented contrast for instance). Other features, and combinations are attested, see Harbour (2014) for discussion.

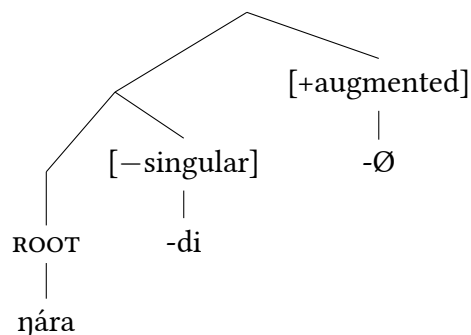
4.3.2 Containment and Markedness

Taking overt encoding of morphemes as our point of departure, recall from above that Manam was characterised as building the dual from the plural:

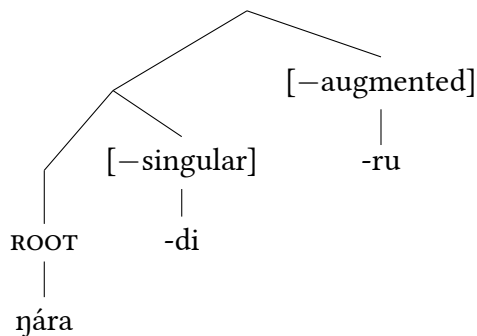
- (28) a. áine **ɲára** b. áine **ɲára-di** c. áine **ɲara-dí-a-ru**
 woman that-SG woman that-PL woman that-PL-LINKER-DL
 ‘that woman’ ‘those women’ ‘those two women’

With reference to the structure in (33), we now understand the containment relation somewhat differently. Manam *-di* is not the plural affix, but is the exponent of [−singular] (see also Nevins 2011), a node that is shared by both the plural and the dual. However, of the two values of [±augmented], only one ([−augmented]) is characterised by an overt exponent, as shown in (34). Indeed, in the representations in (34), it is not true, strictly speaking, that the representation of the dual contains that of the plural in Manam. There is a containment relation between the plural and the dual in the sense that both contain the [−singular] exponent *-di*, and that the dual then contains an additional [−augmented] exponent *-ru*. Crucially, however, under the analysis in (34) this results from there only being an overt [−augmented] exponent and no [+augmented] exponent, rather than structural containment. To put it differently, containment here reflects morpho-phonological exponence, rather than (morpho-)syntactic structure (we return to this point below).

- (34) a. Manam plural

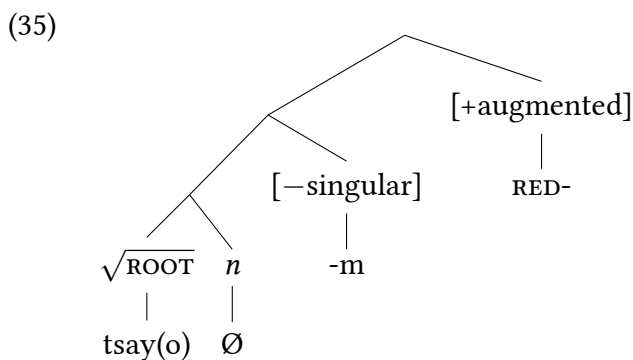


- b. Manam dual



From this perspective, it should come as no surprise that it is possible to preserve the same (morpho-)syntactic (and thus semantic) representations of the dual and plural, but to vary the morpho-phonological relations (overt vs. null) of the exponents of $[\pm\text{augmented}]$. A language in which $[\text{+augmented}]$, rather than $[\text{-augmented}]$ is the sole overtly signalled exponent of that node would then have an apparent containment pattern that is the reverse of Manam. Indeed, this is what is found.

On the view we are considering here, the Hopi word for ‘child’ would have the structure in (35). Abstracting away from the prefixal nature of reduplication (not represented in the tree), this is precisely analogous to Manam (34), except that $[\text{+augmented}]$ is the overtly encoded value for non-singular, rather than $[\text{-augmented}]$, which is \emptyset .



Under our proposal, we do not need to posit that dual contains plural in some languages, with the opposite relation in others: the structure underlying these patterns in our view is always (33), with the cross-linguistic variation lying in which value of $[\pm\text{augmented}]$ receives an overt exponent. In this way, our interpretation of the approach to number put forward by Noyer, Harbour and others, by allowing for variation in overt encoding of the $[\pm\text{augmented}]$ node, provides a succinct characterisation of this variation, while maintaining as invariant the structural representation which underlies Greenberg’s Universal 34.

4.3.3 Suppletion, Markedness and *ABA

The vast majority of suppletive patterns that we have seen for number involve ABB patterns, which are succinctly described as being conditioned by [−singular], the feature that is shared by dual and plural. But since we have now rejected adjacency as a condition on allomorphy, we may describe the suppletion in the root for ‘woman’ in (50) as conditioned by the feature [+augmented], the feature that uniquely characterises plural, correctly capturing the observed pattern. Of course, this raises the spectre of overgeneration. Why could Hopi not just as easily have had a suppletive root triggered by [−augmented]?

We posit that the answer to this question lies in what are possible triggers of suppletion. Here we develop the suggestion of Moskal (2014, 2017), who proposes that there are restrictions concerning markedness on which features can govern suppletion. On the basis of a survey into suppletion patterns found in the inclusive/exclusive distinction in first person pronouns, Moskal concludes that it can be the case that either marked features, or both unmarked and marked features can govern suppletion, but crucially that unmarked features alone are not able to govern suppletion (building on work by Calabrese 2005, Nevins 2010 in phonology). For the clusivity distinction, this means that possible suppletive forms are where the inclusive form is suppletive, compared to the 1SG pronoun, as shown by Evenki in Table 55 (for evidence that the *i~u* alternation marks number, see Table 29 above), or both the inclusive and exclusive pronouns are suppletive, as is seen in Paraguayan Guaraní. Impossible is a language where the inclusive pronoun is non-suppletive but the exclusive is, which is confirmed by the sample investigated in Moskal (2017).⁵⁷

Table 55: Evenki (Nedjalkov 1997)

	SG	PL
1	bi	
1EXCL		bu
1INCL		mit

Table 56: Paraguayan Guaraní (Gregores & Suárez 1967)

	SG	PL
1	še	
1EXCL		ore
1INCL		yane

Here, we propose that markedness correlates with overt encoding. Specifically, if

⁵⁷Moskal (2017) notes that within the realm of clusivity there is variation as to whether the inclusive or the exclusive serves as the base for the other. That is, in some languages, the inclusive form seems to contain the exclusive form, whereas in others, the exclusive form contains the inclusive form. This is the same situation that we note for the containment of dual and plural above. However, Moskal (2017) shows that there is this time no evidence that suppletion also varies along these lines. That is, although containment relations at times suggest the triple SINGULAR—INCLUSIVE—EXCLUSIVE, suppletion patterns never follow this triple. This difference to number goes beyond the scope of our paper, and we refer the reader to Moskal (2017) for further discussion.

$[\alpha\text{AUG}]$ is the marked value, then (i) $[\alpha\text{AUG}]$ is overtly coded and (ii) $[\alpha\text{AUG}]$ can serve as a context for suppletion. Additionally, if $[-\alpha\text{AUG}]$ is unmarked, then (i) $[-\alpha\text{AUG}]$ is phonologically null and (ii) $[-\alpha\text{AUG}]$ cannot serve as a context for suppletion.

In the case of number, we have argued on the basis of the overt morphology, independent of suppletion, that languages vary in which value of $[\pm\text{augmented}]$ is overtly encoded in the context of $[-\text{singular}]$. Marrying markedness and overt encoding then means that in languages that (descriptively) build duals from plurals (like Manam and Sursurunga), $[-\text{augmented}]$ is marked in the context of $[-\text{singular}]$ (cf. Nevins 2011). Conversely, in languages like Mokilese and Panytyima, it is $[\text{+augmented}]$ that is marked, yielding the appearance that plurals are built from duals. Now recall that Hopi, unlike Manam, shows transparent evidence from overt containment morphology (54) that the plural (not the dual) is the marked value among the non-singulars. From the corresponding structure in (35), we predict that *ABA should be read against the sequence SINGULAR-DUAL-PLURAL. Nothing then excludes the plural from suppletting on its own, since $[\text{+augmented}]$ is then the marked value, and can therefore serve as a trigger for suppletion. It is the dual that cannot be the odd member of the paradigm. And this is exactly what we found, not only for Hopi, but for the other three cases of nominal suppletion for number:⁵⁸

Table 57: Lexical noun suppletion.

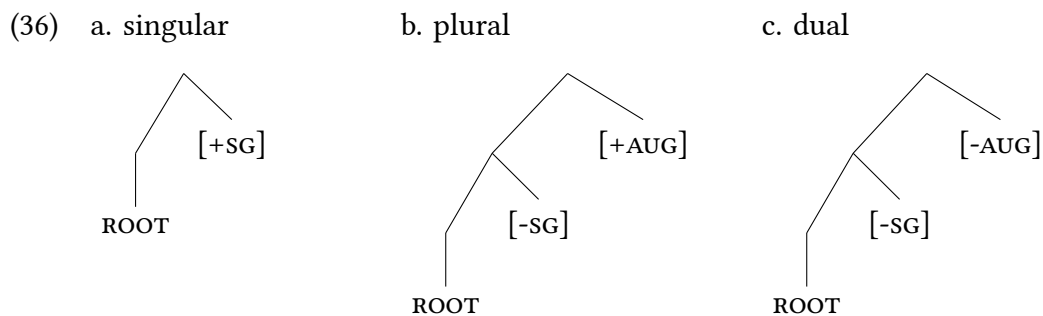
Language	SG	DL	PL	Gloss
Hopi	wùuti	wùutit	momoyam	‘woman’
Lavukaleve	vo’vou	vo’voul	tulav	‘boy’
Slovenian	člòvek	člòvek-a	ljudj-e	‘person’
Yimas	panmal	panmalc-rm	pay-um	‘man’

The initially problematic cases, then, are in fact consistent with the predictions of the theoretical approach, given a more refined understanding of the structural representation of number. By allowing variation in markedness, we allow concomitant variation in suppletive patterns. What we continue to exclude is conflicting patterns: where the suppletive evidence and evidence from overt encoding go in opposite ways. In addition, on the assumption that the representation of number should be consistent within a given language, at least within a single domain (such as nouns), we do not expect variation in suppletive patterns within such a domain.

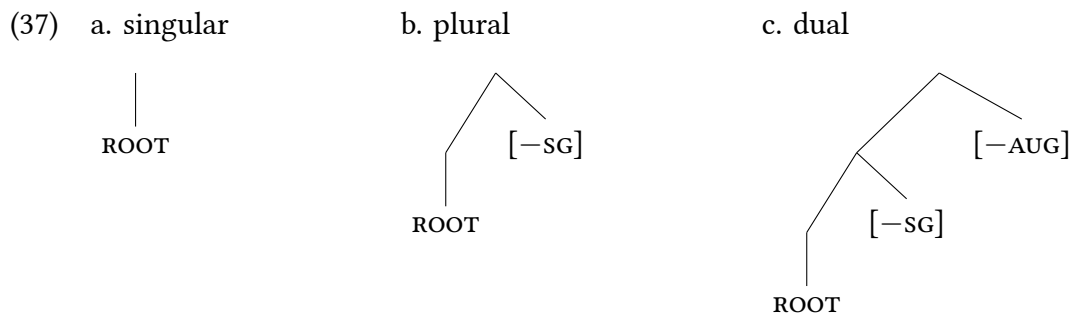
Having established the relation between markedness and overt encoding leaves us with one final issue. At this point, we derive *ABA in two different ways: In adjectives

⁵⁸In the absence of concrete evidence to suggest otherwise, we assume that these show the same markedness ‘reversal’ that Hopi does. Note that there is suggestive evidence that this is the case in Lavukaleve, at least for the case that is listed in Table 57, see footnote 51 above.

and pronominal case it is derived through structural containment, but in pronominal number it is derived through markedness. In a final step, following the spirit of [Calabrese \(2005\)](#), [Nevins \(2010\)](#) and [Moskal \(2014\)](#) where languages vary parametrically as to whether all or only marked features are visible, we propose that markedness has a restrictive effect on Vocabulary Insertion. That is, in some languages Vocabulary Items only have access to marked features. Without the assumption that only marked features are visible, the representations we now assume for singular, dual, and plural are as in (36).⁵⁹



On this representation, there is no literal containment among the various numbers. But if we represent only marked values, then for a language in which ‘dual’ is marked, the relevant representations are as in (37).⁶⁰ Proper containment reemerges and likewise for languages in which plural is marked relative to dual.



Thus, we arrive at the same type of structural containment relation as in adjectival patterns and pronominal case, however, in the case of number this structural containment is derived by a relativisation to marked values.

⁵⁹We omit the [\pm augmented] node in the singular, as the value is redundant, but adding it in (36a) would not affect the point here.

⁶⁰That [+SG] is unmarked, relative to [-SG] in the sense used here is well established: if one value of number is systematically null, with the other value(s) bearing an overt mark, then it is singular which is systematically null ([Corbett 2000](#)). We put aside the interesting question here of the relation of morphological markedness to semantic markedness (on which see [Bobaljik et al. 2011](#)).

A final case to discuss in this vein is Slovenian, which turns out to be the exception that proves the rule. Slovenian has been reported to show exactly the kind of conflicting patterns of suppletion that we do not expect, with the dual of a single noun patterning with singular in some cases, and with the plural in others (the following data are from [Priestly 1993](#) and [Corbett 2007](#)):

Table 58: Slovenian lexical nouns.

	SG	DL	PL
NOMINATIVE	človek	človeka	ljudje
ACCUSATIVE	človeka	človeka	ljudi
GENITIVE	človeka	ljudi	ljudi
DATIVE	človeku	človekoma	ljudem
INSTRUMENTAL	človekom	človekoma	ljudmi
LOCATIVE	človeku	ljudeh	ljudeh

Even the Slovenian data however do not show clear evidence for mixed patterns of suppletion in lexical nouns, once we recognise, as in the discussion of case, the important distinction between syncretism and shared roots. Across the language as a whole, the contrast between dual and plural is neutralised in the genitive and locative cases. These cases show only a singular vs. non-singular contrast – unlike the dual dative or nominative, there is no distinct dual genitive or locative form which shares a root with the plural - these cases simply lack a dual number.

In sum, suppletion for number in pronouns follows the expected pattern if the category of number is internally complex, and if there are containment relations among the values. The ABB and ABC patterns are attested, while ABA is not. The core theoretical prediction is robustly supported. Matters become more complex when we incorporate variation in morpheme order as indicative of containment relations, and when we look at suppletion for number in lexical nouns. As it happens, these two sources of apparently challenging variation can be treated in the same way, once we are more careful with the theory of number and its structural manifestation.

5 Conclusions

In this paper, we have investigated suppletive patterns in case and number in pronouns, an area in which sufficient data is available from a large enough sample of languages to distinguish systematic patterns from accidental gaps. With respect to both case and number we find robust patterns and systematic gaps, mirroring to a large extent the findings in [Bobaljik \(2012\)](#) regarding adjectival suppletion. Most importantly, we find that the ABA pattern of suppletion is unattested in these domains, as it was in

Bobaljik’s survey of adjectival suppletion. A point of difference between adjectives and pronouns is the attestation of AAB patterns in the latter, a difference we take to be informative about the nature of locality between the trigger and target of suppletion. By combining case and number, we find reason to challenge previous proposals that the trigger must be structurally Bobaljik (2012), Adger et al. (2003) or linearly Embick (2010) adjacent to the target (cf. Merchant 2015, Moskal & Smith 2016). We extended Bobaljik’s reasoning about *ABA generalisations to the pronominal domain, and concluded not only that pronouns have internal structure, as is now often argued, but also that both case and number are categories with internal structure. The containment approach to *ABA extends neatly to case on the assumption that oblique cases contain the dependent case, which in turn contain the unmarked case. The suppletion facts and transparent containment patterns track one another as expected. In the domain of number, we find that some notion of markedness plays a role — acknowledging that there is variation in whether the plural or dual is the marked member of the non-singular opposition (at least in the sense of bearing an overt mark), the patterns of suppletion can be characterised as following the logic of containment if one focuses on representations that include only the marked number values.

Perhaps more importantly, our results contribute to a growing body of evidence that finds limits on cross-linguistic variation in large samples. Even suppletion, that most unruly of grammatical phenomena, turns out to be rule-governed when viewed at only a slight level of abstraction. We have argued that simple accounts of the observed limits on variation may be given in structural terms, and in particular, we hope to have demonstrated here that the key ingredients of these accounts extend beyond the phenomena for which they were first posited, providing evidence for general, universal conditions on grammatical representations.

As somewhat of an epilogue, we return to the observation that throughout this study, we have assumed that there is a line to be drawn (even if it is not always clear for the researcher) between suppletion (multiple listed exponents of a morpheme), and “mere” irregularity of other sorts (a single listed exponent, but subject to surface allomorphy via phonological rules). As recognized at various points above, we thus make recourse to morphophonological rules (i.e., morphologically conditioned, phonological rules, such as palatalisation in Polish or diphthongisation in Nepali). Some of our segmentations involve a residue. For German, we identified a consistent formative *d*-marking the second person singular (*du*, *dich*, *dir*, also genitive/possessive *dein*). This *d*-occurs in all and only the second person singular forms among the personal pronouns, and the remaining strings in the non-nominative forms occur independently (as in first person singular *m-ich*, *m-ir*, *m-ein*), but the status of the *u* segment in the nominative of the second person singular is left open. We have held, thus far largely implicitly, that the child is sensitive to such regularities, and encodes them in their grammar: that

d- represents 2SG is something we take to be a synchronic fact of German grammar.⁶¹

One could imagine approaching this from the opposite perspective: that the child starts by listing all forms separately (a form of generalised suppletion, as pursued, e.g., in Siddiqi 2009, Haugen & Siddiqi 2013), and segments forms only when there is robust evidence to do so.⁶² Perhaps the *d-* formative in German was historically a distinct morpheme, but synchronically, *du*, *dich*, and *dir* are lexically listed as full forms with no internal decomposition. Viewing the facts from this perspective would vastly increase the prevalence of suppletion in the world's pronominal systems, in particular a significant number of patterns we treat as ABB (German *ich*, *mich*, *mir*) or AAA (German *du*, *dich*, *dir*) would be characterized as ABC patterns. Whether it would change our major conclusion remains to be seen, and hinges crucially on what constitutes "robust" evidence for segmentation, in particular in typically small, closed systems, such as pronominal inflection. The fact that the analogous debate about the English past tense (whether the forms *ring*~*rang* are related by an unproductive rule or simply listed in exactly the same way as *go*~*went*) show no signs of abating after decades of argument suggests that we will not be able to definitively resolve the issue here. But as one final point in favour of the pronominal decomposition view (rather than generalised suppletion), we might suggest that the very robustness of the *ABA generalization, now established across a range of empirical domains, is a fact in and of itself in need of an explanation. That German *mich* and *mir* share an initial *m-*, which the corresponding nominatives lack, is an observation that can be made independently of whether or not one decomposes these pronouns. But failing to decompose the pronouns treats *ich*, *mich*, *mir* and *du*, *dich*, *dir* equivalently as ABC patterns, and provides no obvious means to exclude *ABA patterns. By contrast, decomposing the pronouns (and the categories of case and number), as we have shown, provides an explanation of the observed facts.

⁶¹As a reviewer and others note, one could ask about German *m-ich* whether an alternative segmentation should be considered, in light of nominative *ich*, which would take the *m-* to be an accusative prefix, unique to the first person singular. While acknowledging that the personal pronoun paradigm is a small, closed class, and that the child acquiring German might consider various possible segmentations, there are more parallels speaking in favour of the analysis we have given. Along with the general observation that German nominal inflection is uniquely suffixing, all of the following pairwise proportional analogies support this analysis, where there is no proportional analogy that can be made in the language to support a putative *m-* accusative prefix: *mich:dich::mir:dir*, *mich:mir::dich:dir*, *mich:mein::dich:dein*, *mich:mein::sich:sein* (and so on for inflected forms of the possessive). We assume that some such tallying goes into the weighting of the likelihood of different competing segmentations.

⁶²We thank Martin Haspelmath, in comments on an earlier draft, for pressing us to be clear about this important issue.

A Appendix A: Case

This appendix lists all the languages examined for case suppletion. For each language, we indicate in the second column (>2K) whether the language has more than two cases (apart from genitive and vocative). For these languages, we indicate whether we have identified suppletion for case, and if so in which pronouns. The online appendix provides the full dataset from all of the languages marked “Y” in the second column, i.e., as having enough case distinctions to be relevant to the study at hand.

A.1 Overview

Language	>2K	Suppletion	Source
Abkhaz	N		Chirikba (2003)
Abui	N		Kratochvil (2007)
Afrikaans	N	AB	Donaldson (1980)
Ainu	N		Tamura (2000)
Alamblak	Y	none	Bruce (1984)
Albanian	Y	ABB: 1sg, 3sg.m; ABC: 3sg.f	Newmark (1982)
Amuesha	N		Duff-Tripp (1997)
Arabela	N		Rich (1999)
Araona	N		Pitman (1980)
Archi	Y	AAB: 2sg, 1sg, 1plexcl, 1plincl; ?ABA: 2pl ⁶³	Kibrik & Kodzasov (1990), Brown et al. (2003)
Armenian	Y	ABB: 1sg, 2sg, 2pl	Kozintseva (1995)
Awa Pit	Y	none	Curnow (1997)
Basaa	N	⁶⁴	Hyman (2003)
Basque	Y	ABB: 3sg.prox	Saltarelli et al. (1988)
Bawm	N		Reichle (1981)
Bengali (Chit-tagong)	Y	none	Ucida (1970)
Bilua	N		Obata (2003)
Brahui	Y	ABB: 1sg	Andronov (1980)
Burmese	Y	none	Okell (1969)
Burushaski	Y	ABB: 2sg, 3sg, 3pl ⁶⁵	Berger (1998)

⁶³See section 3.5.

⁶⁴Some subject pronouns in Basaa are suppletive with respect to the ‘independent’ series, which occurs in all other positions, but it is not clear that this is a case-driven alternation, and in any event, Hyman does not provide evidence for a distinction analyzable as more than a two-way distinction in case.

⁶⁵But see notes in online appendix.

Cahuilla	N		Seiler (1977)
Cavineña	Y	none	Guillaume (2008)
Cyauvava	N		Key (1967)
Chalcatongo	N		Macauley (1996)
Mixtec			
Chawchila	N		Newman (1944)
Chuvash	Y	ABB: 1sg, 2sg, 3sg,	Clark (1998)
Comanche	Y	none	Charney (1993)
Dhaasanach	N		Tosco (2001)
Daga	N		Murane (1974)
Dagaare	N		Bodomo (1997)
Dani (Lower Grand Valley)	N		Bromley (1981)
Danish	N	AB	Allan et al. (1995)
Dumi	Y	none	van Driem (1993)
Dutch	N	AB	co-author's native knowledge
Dyirbal	Y	none	Dixon (1972)
Dzongkha	Y	none	van Driem (1998)
English	N	AB	co-authors' native knowledge
Epena Pedee	Y	none	Harms (1994)
Estonian	Y	none	Viitso (1998)
Evenki	Y	ABB: 1sg, 1pl	Nedjalkov (1997)
Faroese	Y	ABB: 1sg; AB(B): 1pl	Thráinsson et al. (2004)
Fijian, Boumaa	N		Dixon (1988)
Finnish	Y	none ⁶⁶	Karlsson (1999)
French	N	AB	Ferrar (1972)
Fur	Y	none	Beaton (1968)
Garawa	Y	ABB: 1sg, 1du.incl, 3sg	Furby & Furby (1977)
Garo	Y	none	Burling (1961)
Gashowu	N		Newman (1944)
Georgian	Y	ABB: 1sg, 3sg, 3pl	Hewitt (1995)
German	Y	ABB: 1sg, 3sg.m, 3pl; AB(B): 1pl, 2pl,	co-authors' knowledge
Gimira	Y	none	Breeze (1990)
Gooniyandi	Y	AAB: 1sg, 3sg	McGregor (1990)
Greek (Modern)	N	AB	Holton et al. (1997)

⁶⁶A possible ABB paradigm in the singular, animate interrogative pronoun 'who.'

Greenlandic (West)	Y	none	Fortescue (1984)
Hamtai (Kapau)	N		Oates & Oates (1968)
Hua	Y	none	Haiman (1980)
Hungarian	N ⁶⁷		Kenesei et al. (1998)
Hunzib	Y	A(A)B: 2sg	van den Berg (1995)
Icelandic	Y	ABB: 1sg; AB(B): 1pl, 2pl	Einarsson (1945)
Ika	N		Frank (1990)
Imonda	N		Seiler (1985)
Iraqw	N		Mous (1993), Nordbustad (1988)
Itelmen	Y	ABB: 2sg	Field notes (Bobaljic), Volodin (1976)
Japanese	Y	none	Kaiser et al. (2001)
Jingulu	Y	AAB: 1sg?, 2sg	Pensalfini (2003)
Kalispel	N		Vogt (1940)
Kannada	Y	none	Sridhar (1990)
Kanuri	N		Cyffer (1998)
Kashmiri	Y	ABB: 3sg.m.rem 1sg AB(B):	Wali & Koul (1997)
Kayardild	Y	ABB: 2sg	Evans (1995)
Kera	N		Ebert (1979)
Ket	Y	none	Werner (1997)
Kewa	N		Franklin (1971)
Khakas	Y	ABB: 3sg	Baskakov (1975), Brown et al. (2003)
Kham	N		Watters (2002)
Khanty	Y	none	Nikolaeva (1999)
Khwe (Modern)	N		Kilian-Hatz (2008)
Kiowa	N		Watkins (1984)
Klon	N		Baird (2008)
Koasati	N		Kimball (1991)
Korean	Y	none	Lee & Ramsey (2000), co-author's native knowledge
Koromfe	N		Rennison (1997)
Koyra Chiini	N	AB	Heath (2008)

⁶⁷Although more case distinctions are recognised in grammatical descriptions, the pronouns mark only a two-way distinction in case (other than the genitive) – the remaining cases are formed adding personal suffixes to postpositions.

Koyraboro Senni	N		Heath (1999)
Krongo	Y	ABB: 1sg, 2sg, 1plincl, 1plexcl, 2pl, 3sg.m	Reh (1985)
Kunama	N		Bender (1996)
Ladakhi	Y	none	Koshal (1979)
Lango	N		Noonan (1992)
Latvian	Y	ABB: 1sg	Mathiassen (1996)
Lavukaleve	N		Terrill (2003)
Lele	N		Frajzyngier (2001)
Lezgian	Y	none	Haspelmath (1993)
Lithuanian	Y	ABB: 1sg	Mathiassen (1996)
Maba	N		Trenga (1947)
Malakmalak	N		Birk (1976)
Malayalam	Y	ABB: 1sg	Asher & Kumari (1997)
Manam	N		Lichtenberk (1983)
Mangarayi	Y	AAB: 2sg	Merlan (1982)
Maori	N		Bauer (1993)
Mapuche	N		Smeets (2008)
Maranungku	N		Tryon (1970)
Marathi	Y	none	Pandharipande (1997)
Maricopa	Y	none	Gordon (1986)
Martuthunira	Y	none	Dench (1995)
Maybrat	N		Dol (2007)
Meithei/Manipuri	Y	none	Bhat & Ningomba (1997)
Mian	N		Fedden (2007)
Mina	Y	ABB: 1sg,3sg,3pl	Frajzyngier et al. (2005)
Misanltla To- tonac	N		
Miwok (S. Sierra)	Y	none	Broadbent (1964)
Mohawk (Akwe- sasne)	N		Bonvillain (1973)
Mongsen Ao	N		Coupe (2007)
Mongolian (Khalkha)	Y	ABB: 1sg	Poppe (1951)
Mordvin (Erzya)	Y	none	Zaicz (1998)
Mosetén	N		Sakel (2004)
Mundari	N		Osada (1992)
Murle	Y	none	Arensen (1982)

Nenets	N ⁶⁸	AB	Salminen (1998)
Newar (Dolakha)	Y	none	Genetti (2007)
Nez Perce	Y	none	Rude (1985)
Ngiyambaa	Y	none	Donaldson (1980)
Nubian (Don- golese)	N		Armbruster (1960)
Nunggubuyu	Y	none	Heath (1984)
Oromo (Harar)	N		Owens (1985)
Pashto	N	AB	Penzl (1955)
Paumari	N		Chapman & Derbyshire (1991)
Pendau	N		Quick (2007)
Pirahã	N		Everett (1986)
Pitjantjatjara	Y	none	Bowe (1990)
Polish	Y	ABB: 1sg, 3sg.m, 3sg.f, 1pl, 3pl.m	Brooks (1975)
Pomo, Eastern	Y	ABB: 1sg, 2sg	McLendon (1975)
Pomo, S-E (Hokan)	Y	ABB: 1sg, 2sg	Moshinsky (1974)
Puyuma	?	none	Teng (2008)
Quechua (Im- babura)	Y	none	Cole (1982)
Rabha	Y	none	Joseph (2007)
Romani (Kalderash)	Y	ABB: 3sg.m, 3sg.f, 3pl	Boretzky (1994)
Russian	Y	ABB: 1sg, 3sg.m, 3sg.f, 1pl, 3pl	Wade (1992)
Saami (Northern)	Y	none	Nickel (1994)
Semelai	N		Kruspe (1999)
Serbian/Croatian	Y	ABB: 1sg, 3sg.m, 3sg.f, 1pl, 3pl.m	Browne & Alt (2004)
Shipibo-Konibo	?	none	Valenzuela (1997)
Sinaugoro	Y	none	Tauberschmidt (1991)
Sinhala	Y	none ⁶⁹	Gair & Paolillo (1997)

⁶⁸Although Nenets has a 7-way case distinction, pronouns only mark a three-way case distinction, one of which is genitive. All three persons in all numbers appear to undergo suppletion, and would be counted as ABB patterns if genitive were included in the case hierarchy. As in Hungarian, local case forms for pronouns are formed with possessive inflections on postpositions.

⁶⁹Gair & Paolillo (1997, 21) do not give a complete paradigm, but they note that 1sg *mamə* has the oblique stem *maa*, but all other personal pronouns inflect regularly.

Somali	N ⁷⁰		Saeed (1999)
Spanish	N	AB	
Suena	N		Wilson (1974)
Supyire	N		Carlson (1994)
Tamashek	N ⁷¹		Heath (2005)
Tamil	Y	ABB: 1sg, 2sg, 1pl.excl, 2pl	Asher & Kumari (1997)
Thai	N		Iwasaki & Ingkaphirom (2005)
Tiwi	N		Osborne (1974)
Tlingit	N		Naish (1979)
Trumai	Y	none	Guirardello (1999)
Tunen	N	AB	Smith (2011), Isaac (2007)
Tunica	N		Haas (1940)
Turkana	N ⁷²		Dimmendaal (1982)
Turkish	Y	none	Kornfilt (1997)
Udihe	Y	ABB: 1sg, 1pl.excl	Nikolaeva & Tolskaya (2001)
Udmurt	Y	none	Winkler (2001)
Ungarinjin	Y	none	Rumsey (1982)
Urarina	N		Olawsky (2006)
Usan	N		Reesink (1987)
Vietnamese	N		Thompson (1987)
Wambaya	N	AB ⁷³	Nordlinger (1998)
Warao (S. Amer)	N	AB: 1sg	Romero-Figueroa (1997)
Wardaman	Y	AAB: 3sg, 3pl;	Merlan (1994)
Wichita	N		Rood (1976)
Wikchamni	N		Newman (1944)
Wintu	Y	none	Seiler (1977)
Yagua	N		Payne & Payne (1990)
Yanyuwa	Y	AAB: 3sg.m ⁷⁴	Kirton (1996)
Yawelmani (Yokuts)	N		Newman (1944)

⁷⁰Somali has patterns that may be described as ABB and ABC if clitics are included; see the online supplemental materials.

⁷¹Clitics draw a richer case distinction than independent pronouns, and show ABB patterns.

⁷²Apart from genitive, pronouns inflect for absolutive versus nominative. There is also a locative for pronouns that is built from the genitive/possessive, which could be analysed as suppletive relative to the absolutive and nominative forms, yielding AAB patterns, which we have not included.

⁷³Although the pronominal system as a whole draws a three-way case-distinction, no individual pronouns distinguish all three cases

⁷⁴Yanuwa makes a distinction between male and masculine.

Yidiny	Y	¬ABA: 1sg	Dixon (2010)
Yimas	N		Foley (1991)
Yukaghir (Kolyma)	Y	none	Maslova (2003)
Yup'ik (Central)	Y	¬ABA: 1sg	Jacobson (1995)
Yurok	Y	AAB: 3sg	Robins (1958)
Zulu	N		Poulos & Msimang (1998)
!Xóõ	N		Traill (1994)

A.2 ABB Patterns

The following table lists plausible cognate triples of pronouns showing the ABB suppletive patterns for case that we have identified. Since absolute numbers are not relevant, as opposed to the distinction between attested and unattested, we have made a number of educated guesses about cognates without making a careful study of each language. Note that only a single illustrative example of each cognate triple is given, with notes on where other languages have cognate forms given in the final column. For example, the Icelandic 1sg forms *ég - mig - mér* have cognates across Indo-European (Russian: *ja - menja - mne*; Latin *ego - mē - mihi*, etc. (see Table 10 in main text), but as these all descend from a common source, only one example is given in the table. Where it appears to us that a pronominal form may not be cognate with all forms in a related language (as in the Albanian nominative *unë*), we have listed such forms as separate entries.

We have titled the case columns as unmarked (=nominative/absolutive), marked 1 and marked 2. While the general orientation is nominative - accusative - dative or absolutive - ergative - dative, where syncretism would obscure the relevant patterns, we have made substitutions. For example, in Armenian, pronouns do not show a nominative vs. accusative distinction, hence the cases here are nominative/accusative - dative - ablative. Likewise, Albanian first and second person singular pronouns do not distinguish accusative and dative, so we have used nominative - accusative/dative - ablative. As noted in the main text, we have avoided genitive pronouns in this study as we have been unable to systematically distinguish genitive case from possessive pronouns in many of our sources.

Language	Pronoun	Cases			Notes
		unmarked	marked 1	marked 2	
Indo-European:					
Icelandic	1sg	ég	mig	mér	cognates widespread in Indo-European
Albanian	1sg	unë	mua	meje	
Armenian (E)	1sg	es	inj	inj(a)nic	

Armenian (E)	2sg	du	k'ez	k'ez(a)nic	
Russian	1pl	my	nas	nam	cognates across Slavic
Armenian (E)	2pl	duk	jez	jez(a)nic	
Albanian	3sg(m)	ai	(a)të	atij	⁷⁵
German	3sg(m)	er	ihn	ihm	⁷⁶
Kashmiri	3sg(m)	su	təm'	təmis	(remote) ⁷⁷
Serbian	3sg(m)	on	nje-ga	nje-mu	cognates across Slavic ⁷⁸
Serbian	3sg(f)	ona	nju	njoj	cognates across Slavic
Serbian	3pl(m)	oni	njih	njima	cognates across Slavic
Romani (Kalderaš)	3sg(m)	vo(v)	les	lés-kə	⁷⁹
Romani (Kalderaš)	3sg(f)	vój	la	lá-kə	
Romani (Kalderaš)	3sg(f)	von	le	lén-gə	
Armenian (E)	emph	ink'e	iren	irenic	
Dravidian:					
Brahui	1sg	ī	kane	kanki	
Tamil	1sg	naan	en	en-akku	also Malayalam
Tamil	1pl	naanga(l)	enga(l)	engaḷukku	⁸⁰
Tamil	2sg	nii	on	on-akku	
Tamil	2pl	niinga(l)	onga(l)	ongaḷukku	
Turkic:					
Chuvash	1sg	epě	mana	mantan	⁸¹

⁷⁵There is evidently a third person pronominal formative *a-*, alternating with demonstrative *k(ë)-*. While the person formative is thus invariant (AAA), the marking of masculine (contrasting with feminine) shows an ABB pattern in the singular (*-i, -të, -tij*), compared to an AAA pattern in the plural.

⁷⁶We tentatively treat this as synchronically suppletive, although historically, they may share a stem.

⁷⁷Note also corresponding feminine forms: *sə – tami – təmis – etc.* Since gender distinctions are lost in the dative and ablative, the feminine forms have not been counted as distinct from the ABB pattern in the masculine series.

⁷⁸Despite the *-n-* in all three cases, we treat the *on~nj(e)-* alternation as suppletive, as the initial *n-* in the non-nominatives, which occurs only after prepositions in most Slavic languages, does not come from the same source as the *-n* in the nominative (Hill 1977). This suppletive root is shared by all third person pronouns, to which morphology indicating number, gender, and case is added. We list the feminine and plural forms separately below, but as they share a base, they are not truly independent datapoints for suppletion. See also the discussion of Polish in the main text.

⁷⁹As in Slavic, the suppletive third person pronominal base is shared across distinct number and gender forms.

⁸⁰The first and second person plural pronouns clearly inherit the suppletive pattern from the first and second person singular, and in this sense are not independent data points. We count them here since, as discussed in section 3.7, contrasting Tamil to Khakas, suppletive patterns in singular pronouns do not automatically carry over to the corresponding plurals. Although Malayalam shares the suppletive pattern in the singular, the corresponding plural in Malayalam is a transparent AAA pattern.

⁸¹The Chuvash suppletive pattern, not shared with other Turkic languages, is similar to the pattern:

Chuvash	2sg	esě	sana	satan	82
Khakas	3sg	ol	anı	ayaa	83
Tungusic:					
Evenki	1sg	bi	mine(-ve)	min-du	84
Evenki	1pl	bu	mune(-ve)	mun-du	85
Mongolic:					
Khalkha	1sg	bi	namaigi	nadada	
Finno-Ugric:					
Finnish	who.sg	kuka	kene-t	ke-tä	
Kartvelian:					
Georgian	1sg	me	čem-	čem-it	86
Georgian	3sg	is/igi	ma-n	ma-sa	(demonstrative) ⁸⁷
Nakh-Dagestanian:					
Hunzib	2sg	mə	dibi	du-	
Archi	2sg	un	was	wa-	88
Archi	1pl.excl	nen	∅-el	la-	
Archi	1pl.incl	nent'u	nen(a)-∅-u	∅-ela-∅-u	
Archi	3sg.refl	inž	žu- / že-	žu- / že-	
Archi	who	k ^w i	ħ-i	ħa-s	
Chukotko-Kamchatkan:					
Itelmen	2sg	kəzza	kn-anke	kn-anxʔal	
Afro-Asiatic:					
Mina	1sg	sə	kú	k(ù)	
Mina	3sg	Ø/a	Ø/u	ŋ(ù)	
Mina	3pl	i	tətətàn	ŋtətàn/ tətè	
Kadugli-Krongo:					
Kadugli	1sg	àʔàn	nkàtí	kàtí	89

b- mVn - mVn- in the Tungusic languages.

⁸²As mentioned in the main text (see also Daniel 2005), the -s- segment in all columns may be treated as a second person formative, implying that this is not truly suppletive.

⁸³Dative is from *an-ya*. Not all Turkic languages show suppletion in the third singular, but Chuvash shows an apparently cognate suppletive triple.

⁸⁴Cognate: Udihe

⁸⁵Cognate: Udihe

⁸⁶The 1st and 2nd person pronouns are syncretic across nominative, ergative and dative as verbal arguments. The 1SG pronoun suppletes for genitives and obliques, having the genitive/oblique stem *čem-*. That stem is listed here as dative (case 2) as well, since postpositions that govern the dative attach to *čem-*, rather than to *me*, for the 1SG (Hewitt 1995, 25).

⁸⁷The same suppletive pattern occurs in the plural, but the forms are syncretic across the non-nominative cases.

⁸⁸For Archi 2SG and 1PL.EXCL, nominative and ergative are syncretic, thus the forms given here are: nominative (=ergative) - dative - oblique stem. See discussion in section 3.5.

⁸⁹For all Kadugli forms except the third person singular, the first column is syncretic among subject,

Kadugli	2sg	ùʔùŋ	nkòtú	kòtú	
Kadugli	1in	àŋŋá	nkàcá	kàcá	
Kadugli	1ex	óow	nkòtíg	kòtíg	
Kadugli	2pl	àakà	nkàtúkwa	kàtúkwa	
Kadugli	3sg(m)	ìʔùŋ	áníŋ	káníŋ	90
Hokan:					
Pomo (SE)	1sg	ʔa	wi-t	wi-tib	91
Pomo (SE)	2sg	ma	ti	ti-tib	
Australian:					
Garawa	1sg	nayu	ŋana	ŋagi-ndu	
Kayardild	1sg	ngada	ngijuwa	ngijin-da	
Kayardild	2sg	nyingka	ngumbaa	ngumban-da	
Garawa	1du.incl	nuŋgala	niya-nja	niya-ndu	
Garawa	3sg	njulu	ŋaŋa-ndu	ŋaŋa-ŋi	92
Isolates:					
Basque	3sg	hau	honek	honi	93
Burushaski	2sg	un	góo	góor	94
Burushaski	3sg	in(é)	ée	éer	
Burushaski	3pl	u(é)	óo	óor	

A.3 ABC Patterns

Language	Pronoun	Cases	Notes
----------	---------	-------	-------

object, and dative forms, the second and third columns are ablative and locative, respectively.

⁹⁰Columns for the third person are object, dative, locative, respectively.

⁹¹The third ‘case’ indicated here is the benefactive form.

⁹²The analysis of the internal structure of pronouns in many Australian languages is more challenging than for many other languages considered here, in part due to limitations in our expertise. Some patterns classed here and below as ABB or AAB may turn out to be AAA (or ABC) if one understands the relevant properties better, but none seem to have plausible analyses as ABA patterns.

⁹³We follow [Brown et al. \(2003\)](#) in treating *hau*~*hon* as suppletive, though nothing critical hinges on this choice. Note that the suppletion is in the singular only; there is no corresponding alternation in the plural.

⁹⁴Forms here are from the pronominal table in [Berger \(1998, 80\)](#); however Berger’s discussion indicates that the apparent ABB patterns are the conflation of two series of pronouns (a long and a short one). While the relation between some long and corresponding short pronouns is not transparent (as in 2sg *un* ~ *góo*), the description seems to indicate that there is a full regular (AAA) series of long forms for each of the pronouns, and thus these ABB patterns are only apparent. [Grune \(1998, 6\)](#) gives regular (AAA) paradigms for 3rd person, while Berger states that the nominative forms given are demonstratives as there are no 3 person full pronouns.

	unmarked	marked 1	marked 2	
Indo-European:				
Albanian 3sg.f	ajo	(a)të	asaj	95
Nakh-Dagestanian:				
Khinalugh 1sg	zi	jä	as(ir)	

A.4 AAB Patterns

Language	Pronoun	Cases			Notes
		unmarked	marked 1	marked 2	
Algic:					
Yurok	3sg	yoʔ, woʔ, yoʔo·t, woʔo·t	yoʔo·t, woʔo·t	weyaʔik	96
Australian:					
Gooniyandi	1sg	nganyi	nganyi-ngga	ngaddagi	97
Gooniyandi	3sg	niyi	niyi-ngga	nhoowoo	98
Jingulu	1sg	ngaya	ngayarni, ngayirni	ngarr-	99
Jingulu	2sg	nyama	nyamarni	ngaank-, ngank-	100
Mangarayi	2sg	ñan̄gi	ñā-n	ñan̄gi	101
Wardaman	3sg	narnaj	narnaj-(j)i	gunga	
Wardaman	3pl	narnaj-bulu	narnaj-bulu-yi	wurrug	
Yanyuwa	3sg.m	alhi	alhinja	ayu	102

A.5 Other Patterns (Analysis unclear, but implausible as ABA)

⁹⁵If the masculine singular is treated as ABB as in note 75, then the feminine would appear to be ABC (-jo, -të, -saj) once the pronominal formative *a-* is factored out.

⁹⁶See also n. 92

⁹⁷Initial *nga-* also occurs in the second person so cannot be treated as the unique formative for 1sg, suggesting an AAB analysis. Alternatively, it may be AAA with some irregularity.

⁹⁸McGregor (1990, 170) notes that the oblique stem *nhoowoo* corresponds only to the 3sg pronoun, and not to the homophonous determiner *niyi* 'that', consistent with analysing this as a suppletive alternation for the pronoun.

⁹⁹Initial *nga-* also occurs in the second and third person singular accusatives, so cannot be treated as the unique formative for 1sg, suggesting an AAB analysis. Alternatively, it may be AAA with some irregularity. Apparently cognate forms occur in Wambaya, but see n.73.

¹⁰⁰Cognate: Wambaya

¹⁰¹The alternation *ñā-*, *ñā-*, *ñan̄-* is similar to that in neighbouring Jingulu, although these languages are not described as related in, for example, Pensalfini (2001).

¹⁰²This pronoun marks masculine class as opposed to female, and is used only by female speakers (Kirtton 1996, 12). The order of the cases presents a possible challenge to Caha's hierarchy; see note in supplemental online materials.

Language	Pronoun	Cases			Notes
		unmarked	marked 1	marked 2	
Australian: Yidiny	1sg	ŋayu	ŋajɪŋ	ŋaɟu:nda ~ ŋanda	¹⁰³

¹⁰³Initial *ŋa-* is common to all first persons across three numbers. The dual and plural pronouns are readily segmented, but the first singular stem is not.

B Appendix B: Number

B.1 Languages Studied

Language	Suppletion	Form	Source
!Xhoo	none		Traill (1994)
Afrikaans	AB		Donaldson (1993)
Akwesansne Mo-hawk	none		Bonvillain (1973)
Aleut	none		Bergsland (1997)
Ambai	ABB	1/2/3	Smith (2011)
Awtuw	ABB/¬ABA	1/2	Smith (2011)
Bāgandji	none		Hercus (1982)
Bardi	ABB	1incl/2/3	Smith (2011)
Basque	AB		de Rijk (2007)
Belait	ABC	1/2	Smith (2011)
Berik	none		Westrum (1988)
Bilua	none		Obata (2003)
Biri	none		Smith (2011)
Boumaa Fijian	ABB	1excl/1incl/2/3	Dixon (1988)
Bukiyip	ABB/ABC	1/2/3m/3f	Smith (2011)
Bunaba	ABB	1excl/2/3	Smith (2011)
Burushaski	AB		Berger (1998)
Camling	none		Smith (2011)
Carib	none		Courtz (2008)
Cavineña	ABB/ABC	1/3prox	Guillaume (2008)
Chepang	none		Smith (2011)
Comanche	none		Charney (1993)
Crow	none		Graczyk (2007)
Dagaare	AB		Bodomo (1997)
Dehu	ABC/ABB/AAB	1excl/1incl/2/3m	Smith (2011), Tryon (1970)
Djamindjung	ABB/ABC	1excl/1incl/2/3	Smith (2011)
Dolakha Newar	none		Genetti (2007)
Dumi	none		van Driem (1993)
Dyirbal	none		Smith (2011)
Dzongha	none		van Driem (1992)
Eastern Pomo	AB	1	McLendon (1975)
Evenki	none		Smith (2011)
Finnish	none		Karlsson (1999)
Flinders Island	ABC, ¬ABA	1incl/2/3	Smith (2011)

Forest Enets	none		Smith (2011)
Gagadu	ABB/ABC	1incl.m/1incl.f/3m/3f	Smith (2011)
Gothic	ABB	1/2	Smith (2011)
Gurinji	none		Smith (2011)
Hawaaian	ABB	1excl/1incl/3	Smith (2011)
Hopi	AB	1	Forchheimer (1953) ¹⁰⁴
Hua	none		Haiman (1980)
I'saka	none		Donohue & Roque (2004)
Ingush	AB	2	Nichols (2011)
Jarawara	none		Kumar (2012)
Jarnango	ABC	3	Smith (2011)
Jaru	none		Smith (2011)
Jehai	ABB/ABC	1incl/2	Smith (2011)
Jingulu	ABB	2	Pensalfini (1997)
Kamas	ABB	2	Smith (2011)
Kannada	none		Smith (2011)
Karadjeri	none		Smith (2011)
Kayardild	ABB	2/3	Evans (1995)
Ket	none		Smith (2011)
Kham	ABB	1/2	Smith (2011)
Khanty	none		Nikolaeva (1999)
Kilivila	none		Smith (2011)
Koasati	none		Kimball (1991)
Koromfe	AB		Smith (2011)
Kuku-Yalanji	ABC	3	Smith (2011)
Kuna, border	none		Smith (2011)
Kunimaipa	ABB	1excl/1incl/2/3	Smith (2011)
Kwamera	ABB	1excl/1incl/2/3	Smith (2011)
Kwaza	none		van der Voort (2004)
Ladakhi	none		Campbell (2000)
Lavukaleve	ABB	1excl/1incl	Terrill (2003)
Lega-Shabunda	ABB	1/3	Smith (2011)
Lele	none		Frajzyngier (2001)
Lezgian	AB		Haspelmath (1993)
Limbu	none		Smith (2011)
Macushi	AB	1excl	Smith (2011)
Malayalam	none		Asher & Kumari (1997)
Manam	ABB	1excl/1incl/3	Lichtenberk (1983)
Mangala	none		Smith (2011)

¹⁰⁴Hopi has a constructed dual in pronouns, see *e.g.* Corbett (2000).

Mangarayi	ABB	2	Merlan (1982)
Maori	ABB	1excl/1incl/3	Smith (2011)
Mapuche	none		Smeets (2008)
Mapudungun	none		Smith (2011)/lib
Marghi	none		Smith (2011)
Martuthunira	ABB	2	Dench (1995)
Maybrat	AB		Waren (2007)
Mina	none		Frajzyngier et al. (2005)
Mlabri	ABB	2	Smith (2011)
Mokilese	ABB	3	Harbour (2014)
Mongsen Ao	ABB/ABC	1excl/1incl	Coupe (2007)
Navajo	none		Smith (2011)
Ngaju	ABB/ABC	1incl/2/3	Smith (2011)
Ngandi	ABB	1incl	Smith (2011)
Ngarla	none		Smith (2011)
Nishnaabemwin	none		Valentine (2001)
Nyamal	\neg ABA ¹⁰⁵	3	Smith (2011)
Nyigina	ABB	1incl/2/3	Smith (2011)
Nyulnyul	ABB	1incl/2/3	Nekes & Worms (2006)
Nywaygi	ABB/ABC	2/3	Smith (2011)
Ona	none		Smith (2011)
Paamese	ABB	1excl	Crowley (1982)
Panytyima	none		Smith (2011)
Pech	none		Smith (2011)
Pileni	ABB	1excl/1incl/3	Smith (2011)
Pitta-Pitta	ABB/ABC	2/3m.near/3m.general/3m.far	Blake (1979)
Puyuma	none		Teng (2008)
Qiang, northern	ABB	1	Smith (2011)
Rabha	none		Joseph (2007)
Rapa Nui	ABB	1excl/1incl	Smith (2011)
Rotuman	ABB	1excl/1incl/3	Smith (2011)
Samoan	ABB	1excl/1incl/3	Mosel & Hovdhaugen (1992)
Santali	ABB/ABC	1excl/1incl	Smith (2011)
Sanumá	none		Smith (2011)
Savosavo	ABC	2/3	Smith (2011)
Semelai	ABB	1excl/1incl	Smith (2011)
Sinaugoro	none		Tauberschmidt (1991)
Sursurunga	ABB	1excl/1incl/2/3	Harbour (2014)
Tamashek	none		Heath (2005)

¹⁰⁵See section 4.2.1

Tangga	ABB	1excl/1incl/2/3	Smith (2011)
Thai	none		Iwasaki & Ingkaphirom (2005)
Tiri	ABB/ABC	1excl/1incl/2	Smith (2011)
Tokelauan	ABB	1excl/3	Smith (2011)
Toqabaqita	ABB	1excl/1incl/2/3	Lichtenberk (2008)
Tuvalaun	ABB	1excl/1incl/3	Smith (2011)
Urarina	AB	2	Olawsky (2006)
Wajarri	¬ABA ¹⁰⁶	3	Smith (2011)
Wambaya	AAB/ABB	1incl/2	Nordlinger (1998)
Warembori	ABB	1excl/1incl/2/3	Smith (2011)
Warrwa	ABB ¹⁰⁷	1excl/2/3	McGregor (1994)
West landic	Green- none		Fortescue (1984)
Wikngenchera	ABC	3	Smith (2011)
Wunambal	ABB	1excl	Smith (2011)
Yagua	ABB/AAB/¬ABA ¹⁰⁸	1excl/2/3	Payne & Payne (1990)
Yanyuwa	none		Bradley & Kirton (1992)
Yawuru	ABB/ABC	1incl/2/3	Smith (2011)
Yimas	ABBC	1/2	Foley (1991)

B.2 ABB Patterns

Below we list the plausible candidates of ABB patterns for number. Once more, as absolute numbers are not relevant, we have made educated guesses regarding what counts as a cognate.

Language	Pron	Numbers			Notes
		singular	plural	dual	
Austro-Asiatic:					
Semelai	1EXCL	ʔəŋ	yeʔen	yε	109
Semelai	1INCL	ʔəŋ	hεʔen	hε	
Mlabri	2	mεh	bah jum/ʃum	bah	
Austronesian					
Kwamera	1EXCL	iou	kimaha	kimrau	110

¹⁰⁶See section 4.2.1

¹⁰⁷Warrwa has a MINIMAL–AUGMENTED–UNIT–AUGEMENTED system.

¹⁰⁸See section 4.2.1

¹⁰⁹Also Jehai

¹¹⁰Cognates: Ambai, Boumaa Fijian, Hawaiian, Manam, Maori, Mokilese, Paamese, Pileni, Rapa Nui, Rotuman, Samoan, Santali, Sursurunga, Tangga, Tiri, Tokelauan, Toqabaqita, Tuvaluan, Warembori

Kwamera	1INCL	iou	kiitaha	krau	111
Kwamera	2	ik	kimiaha	kimirau	112
Kwamera	3	in	iraha	irau	113
Bukiyip					
Bukiyip	3M	énan / nani	omom mami	omom bwiom	
Bukiyip	3F	okok / kwakwi	owo wawi	echech bwiech	
Bunaban					
Bunaba	1EXCL	ngayini	ngiyirriyani	ngiyirriway	
Bunaba	2	nginji	yinggirriyani	yinggirriway	
Bunaba	3	niy	biyirriyani	biyirriway	
Djamindjungan					
Djamindjung	1EXCL	ŋayug	yirri	yirrinji	
Djamindjung	2	nami	gurri	gurrinji	
Djamindjung	3	dji burri	burrinji		
East Papuan					
Lavukaleve	1EXCL	ngai	e	el	
Lavukaleve	1INCL	ngai	me	mel	
Indo-European					
Gothic	1	ik/mik	weis/uns(is)	wit/ugkis	
Gothic	2	þu/þuk	jus/izwis	jut/igqis	
Gunwingguan					
Gagadu	1INCL.M	ngannj	manaada	manaamana	
Gagadu	1INCL.F	ngannj	maneemba	manaandja	
Ngandi	1INCL	njaka	ŋorrkorr	ŋorrkorni	
Mangarayi	2	ŋianggi	rnurla	rnurr	
Gagadu	3M	ngaayu	nowooda	nowoomana	
Mirndi					
Wambaya	2	nyamirniji	girriyani	gurluwani	114
Niger-Congo					
Lega-Shabunda	1	nne	bíswé	íswé	
Lega-Shabunda	3	gwě	bábo	bo	
Nyulnyulan					
Nyigina	1INCL	ŋayu	yarrdju	yarrdjumirri	115

¹¹¹Cognates: Boumaa Fijian, Hawaiian, Manam, Maori, Mokilese, Pileni, Rapa Nui, Rotuman, Samoan, Sursurunga, Tangga, Tiri, Toqabaqita, Tuvaluan, Warembori

¹¹²Cognates: Ambai, Boumaa Fijian, Ngaju, Sursurunga, Tangga, Dehu, Toqabaqita, Warembori

¹¹³Cognates: Ambai, Boumaa Fijian, Hawaiian, Manam, Maori, Mokilese, Ngaju, Pileni, Rotuman, Samoan, Sursurunga, Tangga, Tokelauan, Toqabaqita, Tuvaluan, Warembori

¹¹⁴Cognate: Jingulu

¹¹⁵Cognates: Bardi, Nyulnyul, Yawuru

Nyigina	2	djuwa	gurrga	gurrgamirri	116
Nyigina	3	ginja/yena	yirrga	yirrgamirri	117
Pama-Nyungan					
Kayardild	2	nyingka	kilda	kirra	118
Kayardild	3	niya	bilda	birra	
Sepik-Ramu					
Awtuw	1	wan	nom	nan	
Sino-Tibetan					
Mongsen Ao	1INCL	ní	ísa/íséŋla	inət	
Kham	1	ŋa:	ge:	gin	119
Kham	2	ni:	je:	jin	
Trans-New-Guinea					
Kunimaipa	1EXCL	ne	rei(paró)	reipi	120
Kunimaipa	1INCL	ne	rari(paró)	raripi	
Kunimaipa	2	ni	ari(paró)	aripi	
Kunimaipa	3	pi	paru(paró)	parupi	
Uralic					
Kamas	2	tan/tên	ši?	šiste	
Uto-Aztecan					
Comanche	1INCL	ni	tannĩ	takwĩ	
Comanche	2	inni	mimmi	mikwĩ	121
Warrwa					
Warrwa	1EXCL	ngayu	yaara, yarrin	yaarawili, yarranbili	
Warrwa	2	juwa	kurra	kurrawili, kurrawawili	
Warrwa	3	kinya	yirra	yirrawili	
Wunambal					
Wunambal	1EXCL	ŋaya	nja:rra	nja:rramiya	
Yagua					
Yagua	1EXCL	ray	núúy	nááy	

B.3 ABC Patterns

Language	Pron	Numbers			Notes
		singular	plural	dual	

¹¹⁶Cognates: Bardi, Nyulnyul, Yawuru

¹¹⁷Cognates: Bardi, Nyulnyul

¹¹⁸Cognates: Nywaygi, Pitta-Pitta

¹¹⁹Cognate: Northern Qiang

¹²⁰Kunimaipa distinctions should be read as MINIMAL–AUGMENTED–UNIT AUGMENTED

¹²¹Warrwa distinctions should be read as MINIMAL–AUGMENTED–UNIT AUGMENTED

Austro-Asiatic					
Jehai	2	moh/miʔ/paj	gin	jih	
Austronesian					
Dehu	1EXCL	ini	eəhun(i)	nyiho	122
Beliat	1INCL	kaw/ko(h), sakay'	kitah, nyakitah	beh-debbeh	123
Beliat	2	naw/no(h), ciw'	(s)unyw	beh(-debbeh), sebbeh	
Bilua					
Bilua	2	ngo	me	qe	
Bilua	3.M.SG.DISTAL	vo	se	nioqa	
Bukiyip					
Bukiyip	1	yek	apak	ohwak	
Bukiyip	2	nyak	ipak	bwiepu	
Caviniña					
Cavineña	1	ike	ekwana	yatse	
Djamindjungan					
Djamindjung	1INCL	ŋayug	yurri	mind	
East-Papuan					
Savosavo	2	no	me	pe	
Savosavo	3M	lo	ze(po)	to	
Gunwinguan					
Gagadu	3F	naawu	nowoomba	ngoyoonjdja/nowoondja	
Nyulnyulan					
Yawuru	3	ginjanğa/yona	yerga/gaŋadjono	njambari/gadambari	
Pama-Nyungan					
Pitta-Pitta	3M.NEAR	ŋuwayi	ṭanayi	pulayi	
Pitta-Pitta	3M.GENERAL	ŋuwaka	ṭanaka	pulaka	
Pitta-Pitta	3M.FAR	ŋuw:rri	ṭana:rri	pula:rri	
Sepik-Ramu					
Yimas	1	ama	ipa	paŋkt	124
Sino-Tibetan					
Mongsen Ao	1EXCL	ní	íla/ikhéla	kenet	

B.4 AAB Patterns

Language	Pron	Numbers			Notes
		singular	plural	dual	
Austronesian					

¹²²Cognate: Ngaju, Santali

¹²³Cognate: Tiri

¹²⁴This is an ABBC pattern. Dual has been excluded from the table. The triple is SINGULAR-PLURAL-PAUCAL.

Dehu	3M	angeice	angate	nyido	125
Mirndi					
Wambaya	1INCL	ngawurniji, ngawu	ngurruwani	mirndiyani	
Yagua					
Yagua	2	jiy	jiryéy	sááda	

B.5 Other Patterns (Analysis unclear, but implausible as ABA)

Language	Pron	Numbers			Notes
		singular	plural	dual	
Austro-Asiatic					
Mlabri	1EXCL	ʔoh	ʔah jum, ʔum	ʔah	
East Papuan					
Savosavo	1EXCL	añi	ave	age	
Lavukalave	2	inu	imi	imil	
Gunwinguan					
Ngandi	1EXCL	ŋaya	ŋjerr	ŋjoworni	
Mangarayi					
Mangarayi	1EXCL	ŋaya	ŋirla	ŋirr	
Mangarayi	1INCL	ŋi	ŋarla	ŋarr	
Pama-Nyungan					
Flinders Island	1INCL	ŋayu	ŋalapal	ŋaluntu	126
Flinders Island	2	yuntu	yarra	yupala	
Nyamal	3	palura	thanalu	piyalu	
Nyamal	3	palura	thanalu	piyalu	
Yanyuwa	3M	(m)yiwa, (w)alhi	alu	wula	
Sepik-Ramu					
Awtuw	2	yen	om	an	
Yagua					
Yagua	3	níí	riy	naada	

References

Acharya, Jayaraj (1991) *A Descriptive Grammar of Nepali and an Analyzed Corpus*. Washington, D.C.: Georgetown University Press.

Adger, David, Susana Bejar, & Daniel Harbour (2003) Directionality of Allomorphy:

¹²⁵Cognate: Jingulu

¹²⁶Cognate: Flinders Island, Jarnango, Kuku-Yalanji, Nyawaygi, Wajarri, Wikngenchera

- A Reply to Carstairs-McCarthy. *Transactions of the Philological Society* 101(1): 109–115.
- Alekseev, M. E. (1985) *Voprosy sravnitel'no-istoričeskoj grammatiki lezginiskix jazykov*. Moscow: Nauka.
- Allan, Robin, Philip Holmes, & Tom Lundskær-Nielsen (1995) *Danish: A Comprehensive Grammar*. London: Routledge.
- Andronov, Mikhael S. (2003) *A comparative grammar of the Dravidian languages*. Berlin: Lincom Europa.
- Andronov, Mikhael S. (1980) *The Brahui Language*. Moscow: Nauka.
- Arensen, Jonathan E. (1982) *Murle grammar*. (Occasional Papers in the Study of Sudanese Languages), Juba, Sudan: Summer Institute of Linguistics and University of Juba.
- Armbruster, Charles Herbert (1960) *Dongolese Nubian: A Grammar*. Cambridge: The University Press.
- Asher, Ronald E. (1982) *Tamil*. Amsterdam: North Holland.
- Asher, Ronald E & T. C. Kumari (1997) *Malayalam*. (Descriptive Grammars Series), London and New York: Routledge.
- Baerman, Matthew (2014) Suppletive kin terms paradigms in the languages of New Guinea. *Linguistic Typology* 18(3): 413–448.
- Baird, Louise (2008) *A Grammar of Klon: A non-Austronesian language of Alor*. Canberra: Pacific Linguistics.
- Baker, Mark (1985) The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16(3): 373–415.
- Bal, Bal Krishna (2007) Structure of Nepali Grammar. In *PAN Localization Working Papers 2004-2007*, Lahore, Pakistan: Center for Research in Urdu Language Processing, National University of Computer and Emerging Sciences.
- Barbiers, Sjef (2007) Indefinite numerals one and many and the cause of ordinal suppletion. *Lingua* 117: 859–880.
- Baskakov, N. A., ed. (1975) *Grammatika xakasskogo jazyka*. Moscow: Nauka.
- Bauer, Winifred (1993) *Maori*. London: Routledge.

- Beaton, A. C. (1968) *A Grammar of the Fur Language*. Khartoum: Sudan Research Unit, University of Khartoum.
- Bender, M. Lionel (1996) *Kunama*. München: Lincom Europa.
- Berger, Hermann (1998) *Die Burushaski-Sprache von Hunza und Nager*. Wiesbaden: Harrassowitz.
- Bergsland, Knut (1997) *Aleut Grammar*. University of Alaska Fairbanks: Alaska Native Language Center.
- Bhat, D. N. S. & M. S. Ningomba (1997) *Manipuri Grammar*. München: Lincom Europa.
- Birk, D. B. W. (1976) *The Malakmalak Language, Daly River (Western Arnhem Land)*. Canberra: Australian National University.
- Blake, Barry J. (1979) Pitta-Pitta. In *Handbook of Australian Languages*, vol. 1, Robert M. W. Dixon & Barry J. Blake, eds., Canberra: Australian National University Press, 182–242.
- Blake, Barry J. (1994) *Case*. Cambridge: Cambridge University Press.
- Bobaljik, Jonathan D. (2000) The Ins and Outs of Contextual Allomorphy. *University of Maryland Working Papers in Linguistics* 10: 35–71.
- Bobaljik, Jonathan D. (2012) *Universals in Comparative Morphology*. Cambridge, MA: MIT Press.
- Bobaljik, Jonathan D. & Heidi Harley (2017) Suppletion is local: Evidence from Hiaki. In *The Structure of Words at the Interfaces*, Heather Newell, Máire Noonan, G L Piggott, & Lisa Travis, eds., Oxford: Oxford University Press.
- Bobaljik, Jonathan D., Andrew Nevins, & Uli Sauerland (2011) Preface: on the morphosemantics of agreement features. *Morphology* 21(2): 131–140.
- Bobaljik, Jonathan D. & Susanne Wurmbrand (2013) Suspension across domains. In *Distributed Morphology today: Morphemes for Morris*, Cambridge, Massachusetts: MIT Press.
- Bobaljik, Jonathan D. & Susi Wurmbrand (2005) The Domain of Agreement. *Natural Language and Linguistic Theory* 23(4): 809–865.
- Bobaljik, Jonathan David & Uli Sauerland (2018) *ABA and the combinatorics of morphological rules. *Glossa* 3(1): 15.
- Bodomo, Adams (1997) *The Structure of Dagaare*. Stanford: CSLI.

- Bonvillain, Nancy (1973) *A Grammar of Akwesasne Mohawk*.
- Boretzky, Norbert (1994) *Romani: Grammatik des Kalderaš-Dialekts mit Texten und Glossar*. Wiesbaden: Harrassowitz Verlag.
- Bowe, Heather (1990) *Categories, Constituents, and Constituent Order in Pitjantjatjara, an Aboriginal Language of Australia*. London: Routledge.
- Bradley, John & Jean Kirton (1992) *Yanyuwa Wuka: A language from Yanyuwa country*. Unpublished manuscript.
- Breeze, Mary J. (1990) *A Sketch of the Phonology and Grammar of Gimira (Benchnon)*. London: School of Oriental and African Studies, University of London.
- Broadbent, Sylvia M. (1964) *The Southern Sierra Miwok Language*. Berkeley: University of California Press.
- Bromley, H. Myron (1981) *A Grammar of Lower Grand Valley Dani*. Canberra: Australian National University.
- Brooks, Maria Z. (1975) *Polish Reference Grammar*. Berlin: de Gruyter.
- Brown, Dunstan, Marina Chumakina, Greville Corbett, & Andrew Hippisley (2003) *The Surrey Suppletion Database*. University of Surrey. <http://dx.doi.org/10.15126/SMG.12/1>.
- Browne, Wayles & Theresa Alt (2004) *A handbook of Bosnian, Serbian, and Croatian*. Durham, North Carolina: Slavic and East European Language Research Center (SEELRC), Duke University.
- Bruce, Les (1984) *The Alamblak Language of Papua New Guinea (East Sepik)*. Canberra: Pacific Linguistics.
- Burling, Robbins (1961) *A Garo Grammar*. Pune: Deccan College Postgraduate and Research Institute.
- Bybee, Joan L. (1985) *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.
- Bye, Patrik (2007) Allomorphy - selection, not optimization. In *Freedom of analysis?*, Sylvia Blaho, Patrik Bye, & Martin Krämer, eds., Berlin: Mouton de Gruyter, 63–92.
- Caha, Pavel (2009) *The nanosyntax of case*. Ph.D. thesis, University of Tromsø.
- Calabrese, Andrea (2005) *Markedness and Economy in a Derivational Model of Phonology*. Berlin: Mouton de Gruyter.

- Calabrese, Andrea (2008) On absolute and contextual syncretism: Remarks on the structure of case paradigms and how to derive them. In *Inflectional Identity*, Asaf Bachrach & Andrew Nevins, eds., Oxford Studies in Theoretical Linguistics, Oxford: Oxford University Press, 156 – 205.
- Campbell, George L. (2000) *Compendium of the World's Languages*, vol. 2: Ladakhi to Zuni. London: Routledge.
- Carlson, Robert (1994) *A Grammar of Supyire*. Berlin: Mouton de Gruyter.
- Chapman, Shirley & Desmond Derbyshire (1991) Paumari. In *Handbook of Amazonian Languages vol.3*, Desmond C. Derbyshire & Geoffrey K. Pullum, eds., Berlin: Mouton de Gruyter, 161–352.
- Charney, Jean Ormsbee (1993) *A Grammar of Comanche*. Lincoln: University of Nebraska Press.
- Chirikba, Viacheslav (2003) *Abkhaz*. München: Lincom Europa.
- Clark, Larry (1998) Chuvash. In *The Turkic Languages*, Lars Johansen & Eva A. Csató, eds., London: Routledge, 434–452.
- Cole, Peter (1982) *Imbabura Quechua*. Amsterdam: North-Holland.
- Corbett, Greville (2000) *Number*. Cambridge: Cambridge University Press.
- Corbett, Greville (2005) Suppletion in personal pronouns: Theory versus practice, and the place of reproducibility in typology. *Linguistic Typology* 9: 1–23.
- Corbett, Greville (2007) Canonical typology, suppletion and possible words. *Language* 83(1): 8–42.
- Coupe, Alexander Robertson (2007) *A Grammar of Mongsen Ao*. Berlin: Mouton de Gruyter.
- Courtz, Henk (2008) *A Carib grammar and dictionary*. Toronto: Magoria Books.
- Crowley, Terry (1982) *The Paamese Language of Vanuatu*. Canberra: Pacific Linguistics.
- Curnow, Timothy J (1997) *A Grammar of Awa Pit (Cuaiquer): An Indigenous Language of south-western Colombia*. Ph.D. thesis, Australian National University, Canberra, Australia.
- Cyffer, Norbert (1998) *A Sketch of Kanuri*. Köln: Rüdiger Köppe.

- Daniel, Michael (2005) Plurality in independent personal pronouns. In *The world atlas of language structures*, Martin Haspelmath, Matthew S. Dryer, David Gil, & Bernard Comrie, eds., chap. 35, Oxford: Oxford University Press.
- de Rijk, Rudolf P. G. (2007) *Standard Basque: A progressive grammar*. Cambridge, MA: MIT Press.
- Dench, Alan Charles (1995) *Martuthunira: A Language of the Pilbara Region of Western Australia*. Canberra: Australian National University.
- Dimmendaal, Gerrit Jan (1982) *The Turkana Language*. Leiden: Rijksuniversiteit te Leiden.
- Dixon, R. M. W. (1972) *The Dyirbal Language of North Queensland. Cambridge Studies in Linguistics*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. (1988) *A Grammar of Boumaa Fijian*. Chicago: University of Chicago.
- Dixon, R. M. W. (1989) *Australian Languages: Their Nature and Development*. Cambridge: Cambridge University Press.
- Dixon, R. M. W. (2010) *A Grammar of Yidij*. Cambridge: Cambridge University Press.
- Dol, Philomena (2007) *A Grammar of Maybrat*. Canberra: Pacific Linguistics.
- Donaldson, Bruce C. (1993) *A Grammar of Afrikaans*. Berlin: Mouton de Gruyter.
- Donaldson, Tamsin (1980) *Ngiyambaa: The Language of the Wangaaybuwan*. Cambridge: Cambridge University Press.
- Donohue, Mark & Lila San Roque (2004) *I'saka*. Canberra: Pacific Linguistics.
- Douglas, Wilfred H. (1981) Watjarri. In *Handbook of Australian Languages: Vol 2*, R. M. W. Dixon & Barry J. Blake, eds., Australian National University.
- Duff-Tripp, Martha (1997) *Gramática del idioma Yanasha' (Amuesha)*. Lima: Ministerio de Educación, Instituto Lingüístico de Verano.
- Ebert, Karen Heide (1979) *Sprache und Tradition der Kera (Tschad). Volume 3: Grammatik*. Berlin: Reimer.
- Einarsson, Stefán (1945) *Icelandic*. Baltimore: Johns Hopkins Press.
- Embick, David (2010) *Localism versus Globalism in Morphology and Phonology*. Cambridge, MA: MIT Press.

- Embick, David & Alec Marantz (2008) Architecture and Blocking. *Linguistic Inquiry* 39(1): 1–53.
- Erschler, David (2017) On multiple exponence, feature stacking, and locality in morphology. Manuscript, University of Massachusetts Amherst.
- Evans, Nicholas (2015) Inflection in Nen. In *The Oxford Handbook of Inflection*, Matthew Baerman, ed., Oxford: Oxford University Press, 545–575.
- Evans, Nicholas D. (1995) *A Grammar of Kayardild*. Berlin: Mouton de Gruyter.
- Everett, Daniel L (1986) Pirahã. In *Handbook of Amazonian Languages 1*, Desmond C. Derbyshire & Geoffrey K. Pullum, eds., Berlin: Mouton de Gruyter, 200–325.
- Fedden, Olcher Sebastian (2007) *A Grammar of Mian*. Ph.D. thesis, University of Melbourne.
- Ferrar, H. (1972) *A French reference grammar*. Oxford: Oxford University Press.
- Foley, William A. (1991) *The Yimas Language of Papua New Guinea*. Stanford: Stanford University Press.
- Forchheimer, Paul (1953) *The Category of Person in Language*. Berlin: Walter de Gruyter.
- Fortescue, Michael (1984) *West Greenlandic*. London: Croom Helm.
- Frajzyngier, Zygmunt (2001) *A Grammar of Lele*. Stanford: CSLI.
- Frajzyngier, Zygmunt, Eric Johnston, & Adrian Edwards (2005) *A Grammar of Mina*. Berlin: Mouton de Gruyter.
- Frank, Paul (1990) *Ika Syntax*. Arlington: Summer Institute of Linguistics and The University of Texas at Arlington.
- Franklin, Karl James (1971) *A Grammar of Kewa, New Guinea*. Canberra: Australian National University.
- Furby, Edward S. & Christine E. Furby (1977) *A preliminary analysis of Garawa phrases and clauses*. Canberra: Pacific Linguistics.
- Gair, James W. & John C. Paolillo (1997) *Sinhala*. München: Lincom Europa.
- Genetti, Carol (2007) *A grammar of Dolakha Newar*. Berlin, New York: Walter de Gruyter.
- Gordon, Lynn (1986) *Maricopa Morphology and Syntax*. Berkeley: University of California Press.

- Graczyk, Randolph (2007) *A Grammar of Crow*. Lincoln, Nebraska: University of Nebraska Press.
- Greenberg, Joseph H. (1963) Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of language*, Joseph H. Greenberg, ed., Cambridge, MA: MIT Press, 73–113.
- Gregores, Emma & Jorge Alberto Suárez (1967) *A description of colloquial Guaraní*. Mouton & Company.
- Grune, Dick (1998) *Burushaski: An Extraordinary Language in the Karakoram Mountains*. Pontypridd, Wales: Joseph Biddulph.
- Guillaume, Antoine (2008) *A Grammar of Cavineña*. Berlin: Mouton de Gruyter.
- Guirardello, Raquel (1999) *A Reference Grammar of Trumai*. Houston, Tex: Rice University Press.
- Haas, Mary R. (1940) *Tunica*. New York: J. J. Augustin.
- Haiman, John (1980) *Hua: A Papuan Language of the Eastern Highlands of New Guinea*. Amsterdam: John Benjamins.
- Halle, Morris & Alec Marantz (1993) Distributed Morphology and the Pieces of Inflection. In *The View from Building 20*, Ken Hale & Samuel Jay Keyser, eds., Cambridge, MA: MIT Press, 111–176.
- Harðarson, Gísli Rúnar (2016) A case for a Weak Case Contiguity hypothesis—a reply to Caha. *Natural Language & Linguistic Theory* 34(4): 1329–1343.
- Harbour, Daniel (2008) *Morphosemantic number: From Kiowa noun classes to UG number features*. New York: Springer.
- Harbour, Daniel (2011) Valence and Atomic Number. *Linguistic Inquiry* 42(4): 561–594.
- Harbour, Daniel (2014) Paucity, abundance, and the theory of number. *Language* 90(1): 185–229.
- Harbour, Daniel (2017) Frankenduals: Their typology, structure, and significance. Unpublished manuscript, Queen Mary University of London.
- Harms, Philip Lee (1994) *Epena Pedee Syntax*. Dallas / Arlington: Summer Institute of Linguistics and University of Texas.
- Haspelmath, Martin (1993) *A Grammar of Lezgian*. Berlin: Mouton de Gruyter.

- Haugen, Jason D. & Daniel Siddiqi (2013) Roots and the derivation. *Linguistic Inquiry* 44(3): 493–517.
- Heath, Jeffrey (1984) *A Functional Grammar of Nunggubuyu*. Atlantic Highlands N. J. / Canberra: Humanities Press / Australian Institute of Aboriginal Studies.
- Heath, Jeffrey (1999) *A Grammar of Koyraboro (Koroboro) Senni: The Songhay of Gao, Mali*. Köln: R. Köppe.
- Heath, Jeffrey (2005) *A Grammar of Tamashek*. Berlin: Mouton de Gruyter.
- Heath, Jeffrey (2008) *A Grammar of Koyra Chiini: The Songhay of Timbuktu*. Berlin: de Gruyter.
- Hercus, Luise (1982) *The Bāgandji Language*. Pacific Linguistics 67, Canberra, Australia: Australian National University.
- Hewitt, B. G. (1995) *Georgian: A Structural Reference Grammar*. Amsterdam: John Benjamins.
- Hill, Steven P. (1977) *The N-factor and Russian prepositions*. The Hague: Mouton.
- Holton, David, Peter Mackridge, & Irene Philippaki-Warbuton (1997) *Greek: A comprehensive grammar of the modern language*. London: Routledge.
- Hyman, Larry (2003) Basaa (A.43). In *The Bantu Languages*, Derek Nurse & Gérard Philippson, eds., London: Routledge, 257–282.
- Ibragimov, Garun (1978) *Rutul'skij jazyk. [Rutul]*. Moscow: Nauka.
- Isaac, Kendall Mark (2007) *Participant reference in Tunen narrative discourse*. MA thesis, Graduate Institute of Applied Linguistics, Dallas, Tex.
- Iwasaki, Shoichi & Preeya Ingkaphirom (2005) *A reference grammar of Thai*. Cambridge: Cambridge University Press.
- Jacobson, Steven A. (1995) *A Practical Grammar of the Central Alaskan Yup'ik Eskimo Language*. Fairbanks: Alaska Native Language Center.
- Jakobson, Roman (1936/1971) Beitrag zur allgemeinen Kasuslehre. Gesamtbedeutungen der russischen Kasus. In *Selected Writings*, vol. 2, The Hague: Mouton, 23–71.
- Joseph, Umbavu Varghese (2007) *Rabha*. Leiden: Brill.
- Kaiser, Stefan, Yasuko Ichikawa, Noriko Kobayashi, & Hilofumi Yamamoto (2001) *Japanese: A Comprehensive Grammar*. London / New York: Routledge.

- Karlsson, Fred (1999) *Finnish: An essential grammar*. London: Routledge.
- Katz, Joshua (1998) *Topics in Indo-European Personal Pronouns*. Ph.D. thesis, Harvard University.
- Kayne, Richard (2005) *Movement and silence*. Oxford: Oxford University Press.
- Kenesei, István, Robert M. Vago, & Anna Fenyvesi (1998) *Hungarian*. London: Routledge.
- Key, Harold H. (1967) *Morphology of Cayuvava*. The Hague: Mouton.
- Kibrik, A. E. & S. V. Kodzasov (1990) *Sopostavitel'noe izučenie dagestanskix jazykov. Imja. Fonetika*. Moscow: Moscow University Press.
- Kilian-Hatz, Christa (2008) *A Grammar of Modern Khwe (Central Khosian)*. Köln: R. Köppe.
- Kimball, Geoffrey (1991) *Koasati Grammar*. Lincoln, Nebraska: University of Nebraska Press.
- Kiparsky, Paul (1973) "Elsewhere" in phonology. In *A Festschrift for Morris Halle*, New York: Holt, Reinhart and Winston, 93–106.
- Kirton, Jean F. (1996) *Further aspects of the grammar of Yanyuwa*. Canberra: Pacific Linguistics.
- Kornfilt, Jaklin (1997) *Turkish*. London: Routledge.
- Koshal, Sanyukta (1979) *Ladakhi Grammar*. Delhi: Motilal Banarsidasa.
- Kozintseva, Natalia A. (1995) *Modern Eastern Armenian*. München: Lincom Europa.
- Kratochvil, Frantisek (2007) *A Grammar of Abui*. Utrecht: LOT Dissertations in Linguistics.
- Kruspe, Nicole (1999) *A Grammar of Semelai*. Ph.D. thesis, University of Melbourne.
- Kumar, Pramod (2012) *Descriptive and typological study of Jarawa*. Ph.D. thesis, Jawaharlal Nehru University.
- LaPolla, Randy J. & Chenglong Huang (2003) *A grammar of Qiang*. Berlin: Mouton de Gruyter.
- Lee, Iksop & Robert Ramsey (2000) *The Korean Language*. Albany: State University of New York Press.

- Lenormand, Maurice-Henry (1999) *Dictionnaire drehu-français: La langue de Lifou, Iles Loyalty*. Nouméa, New Caledonia: Le Rocher-à-la-voile.
- Lichtenberk, Frantisek (1983) *A Grammar of Manam*. Honolulu: University of Hawaii Press.
- Lichtenberk, František (2008) *A Grammar of Toqabaqita*. Berlin: Mouton de Gruyter.
- Macauley, Monica (1996) *A Grammar of Chalcatongo Mixtec*. Berkeley: University of California Press.
- Marantz, Alec (2000) Words. Manuscript, MIT.
- Marantz, Alec (2007) Phases and words. In *Phases in the theory of grammar*, Sook-Hee Choe, ed., Seoul: Dong In, 199–222.
- Maslova, Elena (2003) *A Grammar of Kolyma Yukaghir*. Berlin; New York: Mouton de Gruyter.
- Mathiassen, Terie (1996) *A Short Grammar of Lithuanian*. Columbus: Slavica Publishers.
- McCreight, Katherine & Catherine V. Chvany (1991) Geometric representation of paradigms in a modular theory of grammar. In *Paradigms: The economy of inflection*, Frans Plank, ed., Berlin: Mouton de Gruyter, 91–112.
- McFadden, Thomas (2014) Why nominative is special: Stem-allomorphy and case structures. Talk given at GLOW 37, Brussels.
- McFadden, Thomas (2018) *ABA in stem-allomorphy and the emptiness of the nominative. *Glossa: a journal of general linguistics* 3(1): 1–36.
- McGregor, William (1990) *A Functional Grammar of Gooniyandi*. Amsterdam: John Benjamins.
- McGregor, William (1994) *Warrwa, Languages of the World/Materials*, vol. 89. München, Newcastle: Lincom Europa.
- McLendon, Sally (1975) *A Grammar of Eastern Pomo*. Berkeley, CA: University of California Press.
- Mel'čuk, Igor A. (1994) Suppletion: Toward a Logical Analysis of the Concept. *Studies in Language* 18(2): 339–410.
- Merchant, Jason (2015) How much context is enough? Two cases of span-conditioned stem allomorphy. *Linguistic Inquiry* 46(2): 273–303.

- Merlan, Francesca C. (1982) *Mangarayi*. Amsterdam: North-Holland.
- Merlan, Francesca C. (1994) *A Grammar of Wardaman, a language of the Northern Territory of Australia*. Berlin: Mouton de Gruyter.
- Mosel, Ulrike & Even Hovdhaugen (1992) *Samoan Reference Grammar*. Oslo: Scandinavian University Press.
- Moshinsky, Julius (1974) *A Grammar of Southeastern Pomo*. Berkeley / Los Angeles, CA: University of California Publications in Linguistics.
- Moskal, Beata (2013) On some suppletion patterns in nouns and pronouns. Talk given at PhonoLAM, Meertens Instituut.
- Moskal, Beata (2014) The role of morphological markedness in exclusive/inclusive pronouns. In *The proceedings of the 40th Berkeley Linguistics Society*, 354–368.
- Moskal, Beata (2015a) *Domains on the border: Between Morphology and Phonology*. Ph.D. thesis, University of Connecticut, Storrs, CT.
- Moskal, Beata (2015b) Limits on allomorphy: A case-study in nominal suppletion. *Linguistic Inquiry* 46(2): 363–375.
- Moskal, Beata (2017) Suppletion patterns in clusivity and the role of markedness, manuscript, Goethe-Universität Frankfurt.
- Moskal, Beata & Peter W. Smith (2016) Towards a theory without adjacency: Hyper-contextual VI-rules. *Morphology* 26(3-4): 295–312.
- Mous, Maarten (1993) *A Grammar of Iraqw*. Hamburg: Helmut Buske Verlag.
- Müller, Gereon (2004) On decomposing inflection class feature: Syncretism in Russian noun inflection. In *Explorations in Nominal Inflection*, G. Müller, L. Gunkel, & G. Zifonum, eds., Berlin: Mouton de Gruyter, 189–227.
- Murane, Elizabeth (1974) *Daga Grammar: From Morpheme to Discourse*. (Summer institute of linguistics publications in linguistics and related fields, no. 43), Norman, Okla.: Summer Institute of Linguistics, University of Oklahoma.
- Naish, Constance (1979) *A Syntactic Study of Tlingit*. Dallas: Summer Institute of Linguistics.
- Nedjalkov, Igor (1997) *Evenki*. London / New York: Routledge.
- Nekes, Hermann & Ernest Worms (2006) *Australian Languages*. Berlin: Mouton de Gruyter.

- Nevins, Andrew (2010) *Locality in Vowel Harmony*. Cambridge, MA: MIT Press.
- Nevins, Andrew (2011) Marked Targets versus Marked Triggers and Impoverishment of the Dual. *Linguistic Inquiry* 42(3): 413–444.
- Newman, Stanley (1944) *The Yokuts Language of California*. New York: The Viking Fund Publications in Anthropology.
- Newmark, Leonhard (1982) *Standard Albanian: A reference grammar for students*. Stanford: Stanford University Press.
- Nichols, Johanna (2011) *Ingush Grammar*. Berkeley, CA: University of California Press.
- Nickel, Klaus Peter (1994) *Samisk grammatikk*. Karasjok: Girji.
- Nikolaeva, Irina (1999) *Ostyak*. München: Lincom Europa.
- Nikolaeva, Irina & Maria Tolskaya (2001) *A Grammar of Udihe*. Berlin: Mouton de Gruyter.
- Noonan, Michael (1992) *A Grammar of Lango*. Berlin: Mouton de Gruyter.
- Nordbustad, Frøydis (1988) *Iraqw Grammar: An Analytical Study of the Iraqw Language*. Berlin: Dietrich Reimer Verlag.
- Nordlinger, Rachel (1998) *A Grammar of Wambaya, Northern Territory (Australia)*. Canberra: Australian National University.
- Noyer, Rolf (1992) *Features, positions and affixes in autonomous morphological structure*. Ph.D. thesis, MIT.
- Oates, William J. & Lynette F. Oates (1968) *Kapau Pedagogical Grammar*. Canberra: Australian National University.
- Obata, Kazuko (2003) *A Grammar of Bilua*. Canberra: Pacific Linguistics.
- Okell, John (1969) *A Reference Grammar of Colloquial Burmese (two volumes)*. London: Oxford University Press.
- Olawsky, Knut (2006) *A Grammar of Urarina*. Berlin: Mouton de Gruyter.
- Osada, Toshiki (1992) *A Reference Grammar of Mundari*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Osborne, C. R. (1974) *The Tiwi Language*. Canberra: Australian Institute of Aboriginal Studies.

- Owens, Jonathan (1985) *A Grammar of Harar Oromo (Northeastern Ethiopia)*. Hamburg: Helmut Buske Verlag.
- Palmer, Bill (2016) *The Languages and Linguistics of the New Guinea Area: A Comprehensive Guide*. Berlin: Walter de Gruyter.
- Pandharipande, Rajeshwari V. (1997) *Marathi*. London: Routledge.
- Pantcheva, Marina (2011) *Decomposing path: The nanosyntax of spatial expressions*. Ph.D. thesis, Universitetet i Tromsø.
- Payne, Doris L. & Thomas Payne (1990) Yagua. In *Handbook of Amazonian Languages* 2, Desmond C. Derbyshire & Geoffrey K. Pullum, eds., Berlin: Mouton de Gruyter, 249–474.
- Pensalfini, Robert (2001) On the typological and genetic affiliation of Jingulu. In *Forty years on: Ken Hale and Australian languages*, Jane Simpson, David Nash, Mary Laughren, Peter Austin, & Barry Alpher, eds., Canberra: Pacific Linguistics, 385–399.
- Pensalfini, Robert J. (1997) *Jingulu grammar, dictionary, and texts*. Ph.D. thesis, MIT, Cambridge, MA.
- Pensalfini, Robert J. (2003) *A grammar of Jingulu: An aboriginal language of the Northern Territory*. Canberra: Pacific Linguistics.
- Penzl, Herbert (1955) *A grammar of Pashto: A descriptive study of the dialect of Kandahar*. Washington: American Council of Learned Societies.
- Pitman, Donald (1980) *Bosquejo de la gramática araona*. Riberalta, Bolivia: Instituto Lingüístico de Verano.
- Polinsky, Maria, Nina Radkevich, & Marina Chumakina (In press) Agreement between arguments? Not really. In *Verbal Domain*, Roberta D'Alessandro, Irene Franco, & Ángel J. Gallego, eds., Oxford: Oxford University Press.
- Poppe, Nikolaus (1951) *Khalkha-Mongolische Grammatik*. Wiesbaden: Franz Steiner Verlag GMBH.
- Poulos, George & Christian T. Msimang (1998) *A Linguistic Analysis of Zulu*. Cape Town: Via Afrika.
- Priestly, T. (1993) Slovene. In *The Slavonic languages*, vol. 1, Bernard Comrie & Greville Corbett, eds., London: Routledge, 388–451.

- Quick, Phil (2007) *A Grammar of the Pendau Language of central Sulawesi, Indonesian*. Canberra, Australia: Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University.
- Radkevich, Nina (2010) *On Location: The Structure of Case and Adpositions*. Ph.D. thesis, University of Connecticut, Storrs, CT.
- Reesink, Ger P. (1987) *Structures and Their Functions in Usan: a Papuan Language of Papua New Guinea*. Amsterdam: John Benjamins.
- Reh, Mechthild (1985) *Die Krongo-Sprache (Ninò Mó-Dì)*. Berlin: Dietrich Reimer Verlag.
- Reichle, Verena (1981) *Bawm Language and Lore. Tibeto-Burman Area*. Bern: Peter Lang.
- Rennison, John R. (1997) *Koromfe*. London: Routledge.
- Rich, Rolland (1999) *Diccionario Arabela-Castellano*. Lima: Instituto Lingüístico De Verano.
- Robins, R. H. (1958) *The Yurok Language, Grammar, Texts, Lexicon*. Berkeley / Los Angeles: University of California Press.
- Romero-Figueroa, A. (1997) *A reference grammar of Warao*. München: Lincom Europa.
- Rood, David S. (1976) *Wichita Grammar*. New York: Garland.
- Rubach, Jerzy (1984) *Cyclic and lexical phonology: The structure of Polish*. Dordrecht: Foris.
- Rude, Noel E. (1985) *Studies in Nez Perce Grammar and Discourse*. Ph.D. thesis, Eugene: University of Oregon.
- Rumsey, A. (1982) *An Intra-Sentence Grammar of Ungarinjin, North-Western Australia*. Canberra: Australian National University.
- Saeed, John I. (1999) *Somali*. Amsterdam and Philadelphia: John Benjamins.
- Sakel, Jeanette (2004) *Mosetén*. Berlin: Mouton de Gruyter.
- Salminen, Tapani (1998) Nenets. In *The Uralic Languages*, Daniel Abondolo, ed., London: Routledge, 516–547.
- Saltarelli, Mario, Miren Azkarate, David Farwell, Jon de Urbina, & Lourdes Oñederra (1988) *Basque*. London: Croom Helm.

- Seiler, Hansjakob (1977) *Cahuilla Grammar*. Banning, CA: Malki Museum Press.
- Seiler, Walter (1985) *Imonda, a Papuan Language*. Canberra: Australian National University.
- Siddiqi, Daniel (2009) *Syntax within the word: Economy, allomorphy and argument selection in Distributed Morphology*. Amsterdam/Philadelphia: John Benjamins.
- Smeets, Inke (2008) *A Grammar of Mapuche*. Berlin: Mouton de Gruyter.
- Smith, Norval (2011) Free Personal Pronoun System Database. Online at <http://languagelink.let.uu.nl/fpps/>.
- Smith, Peter W., Beata Moskal, Ting Xu, Jungmin Kang, & Jonathan D. Bobaljik (2015) Pronominal Suppletion: Case and Number. In *The Proceedings of NELS 45*, vol. 3, Thuy Bui & Deniz Özyıldız, eds., Amherst, MA: GLSA, 69–78.
- Spencer, Andrew & Gregory T. Stump (2013) Hungarian pronominal case and the dichotomy of form in inflectional morphology. *Natural Language and Linguistic Theory* 31: 1207–1248.
- Sridhar, S. N. (1990) *Kannada: Descriptive Grammar*. London: Routledge.
- Tamura, Suzuko (2000) *The Ainu Language*. Tokyo: Sanseido.
- Tauberschmidt, Gerhard (1991) *A Grammar of Sinaugoro*. Canberra: Pacific Linguistics.
- Teng, Stacy Fang-Ching (2008) *A Reference Grammar of Puyuma*. Canberra: Pacific Linguistics.
- Terrill, Angela (1998) *Biri*. München: Lincom Europa.
- Terrill, Angela (2003) *A Grammar of Lavukaleve*. Berlin: Mouton de Gruyter.
- Thomas, David (1955) Three analyses of the Ilocano pronoun system. *Word* 11(2): 204–208.
- Thompson, Laurence C (1987) *A Vietnamese Reference Grammar*. Honolulu: University of Hawaii Press.
- Thráinsson, Hoskuldur, Hjalmar P. Petersen, Jógvan Jacobsen, & Zakaris Svabo Hansen (2004) *Faroese: An Overview and Reference Grammar*. Tórshavn: Foroya Fróðskaparfelag.
- Tosco, Mauro (2001) *The Dhaasanac language: grammar, texts and vocabulary of a Cushitic language of Ethiopia*. Köln: Köppe.

- Trails, Anthony (1994) *A !Xóõ Dictionary*. Köln: Rüdiger Köppe Verlag.
- Trenga, Georges (1947) *Le Bura-Mabang du Ouadaï: Notes pour servir à l'étude de la langue Maba*. Paris: Institut d'Ethnologie, Université de Paris.
- Trommer, Jochen (2008) Case suffixes, postpositions and the phonological word in Hungarian. *Linguistics* 46: 403–438.
- Tryon, Darrell T. (1970) *An Introduction to Maranungku*. Canberra: Australian National University.
- Učida, Norihiko (1970) *Der Bengali-Dialekt von Chittagong: Grammatik, Texte, Wörterbuch*. Wiesbaden: Harrassowitz.
- Valentine, Randy (2001) *Nishnaabemwin Reference Grammar*. Toronto: University of Toronto Press.
- Valenzuela, Pilar M. (1997) *Basic Verbs Types and Argument Structures in Shipibo-Conibo*. Master's thesis, University of Oregon.
- van den Berg, Helma (1995) *A Grammar of Hunzib*. München: Lincom Europa.
- van der Voort, Hein (2004) *A Grammar of Kwaza*. Berlin: Mouton de Gruyter.
- van Driem, George (1992) *The Grammar of Dzongkha*. Dzongkha Development Commission.
- van Driem, George (1993) *A Grammar of Dumi*. Berlin: Mouton de Gruyter.
- van Driem, George (1998) *Languages of the Greater Himalayan Region*. Leiden: Brill.
- Veselinova, Ljuba N. (2006) *Suppletion in verb paradigms*. Amsterdam: John Benjamins.
- Vitso, Tiit-Rein (1998) Estonian. In *The Uralic Languages*, Daniel Abondolo, ed., London / New York: Routledge, 115–148.
- Vogt, Hans (1940) *The Kalispel Language*. Oslo: Norwegian Academy of Sciences / Jacob Dybwad.
- Volodin, Aleksandr P. (1976) *Itel'menskij jazyk*. Leningrad: Nauka.
- Wade, Terence L. B. (1992) *A Comprehensive Russian Grammar*. Oxford: Blackwell.
- Wali, Kashi & Omkar N. Koul (1997) *Kashmiri: A Cognitive-Descriptive Grammar*. London: Routledge.

- Waren, Olivia Ursula (2007) Possessive pronouns in Maybrat: A Papuan language. *Linguistika* 14.
- Watkins, Laurel J (1984) *A Grammar of Kiowa*. Lincoln: University of Nebraska Press.
- Watters, David E. (2002) *A grammar of Kham*. Cambridge: Cambridge University Press.
- Werner, Heinrich (1997) *Die Ketische Sprache*. Wiesbaden: Harrassowitz Verlag.
- Westrum, Peter N. (1988) *A Grammatical Sketch of Berik*, mA Thesis, University of North Dakota.
- Wilson, Darryl (1974) *Suena Grammar*. Workpapers in Papua New Guinea Languages 8, Summer Institute of Linguistics.
- Winkler, Eberhard (2001) *Udmurt*. München: Lincom Europa.
- Zaicz, Gábor (1998) *Mordva*. London: Routledge.
- Zompì, Stanislao (2017) *Case decomposition meets dependent-case theories*. Master's thesis, Università de Pisa.