Learning (im)possible number syncretisms: Investigating featural organization

Word count: 14,997

Abstract Number features deriving singular, dual, and plural values — always referred to in that schematic (sg-du-pl / ABC) order — have been argued to be organized *via cross-classification* (e.g. Harley & Ritter 2002; Harbour 2014), as well as *in linear hierarchies* (e.g. Harley 1994; Smith et al. 2019). Empirically, these approaches split on which pairs of those number values should readily neutralize to a single form: they agree that the dual and plural values (ABB) form such a natural class, but make opposing predictions for the natural class membership of the singular value. This paper tests their predictions experimentally, by teaching participants artificial languages differing in which of those three logically possible pairs of number values participate in syncretism. Results from 149 adult participants show that participants exposed to the dual-plural syncretism (ABB) had the greatest success in a referent selection task. This supports Maldonado & Culbertson's (2020) findings that adults employ *morphological* features in artificial language learning experiments. Participants also readily learned the singular-plural (ABA) syncretism, but had significant difficulty with the singular-dual (AAB) syncretism ($p < 0.001$). This suggests support for a linear hierarchy approach to number features.

Keywords: number, features, cross-classification, linear hierarchy, syncretism, learnability

1 Introduction

The category of number is often represented using features: a small set of primitives whose interactions are meant to generate the number contrasts and reflect the morphological and semantic patterns attested across languages. Take the dual, which marks quantities of two. In singular-dual-plural languages — always referred to in that schematic (sg-du-pl / ABC) order — the dual contrasts two of something against only one (singular), and against any other quantity (plural). But in singular-plural languages like English, quantities of two are always marked as plural, not singular; when the dual is altogether "absent from the syntax of a particular language", its meaning is conflated with the plural (McGinnis 2005: 700–1; Corbett 2000). This is semantically natural once we consider the meaning of the dual — pairs, after all, are nonsingular. The dual and the plural (ABB) are evidently/empirically a morphological and semantic natural class.

Though every number feature theory takes a dual-plural natural class as a necessary outcome, approaches diverge on how featural primitives are *organized* to capture that affiliation. Are number values built from *cross-classifying* subparts, with dual and plural sharing a subpart in common — the way s'mores and New York-style cheesecake both contain graham crackers as an ingredient? Or are number values *linearly dependent* on one another, with one requiring the presence of the other as a proper subpart — the way a lemon meringue pie encompasses a lemon tart? This choice of feature organization dictates what the grammar is able to make direct reference to, as natural classes that can pattern together morphologically. Both organizational schemes pick out dual and plural as one natural class, but each derives a different *additional* natural class. This paper reports on an artificial language learning experiment that tests these predictions.

1.1 Cross-classifying features

Cross-classifying feature theories of number take dual and plural to share a common subpart. That particular shared feature classifies dual and plural into the category of nonsingulars; other cross-cutting features classify singular, dual, and plural into other natural class categories. For example, singletons and pairs are both smallest possible versions of something: they are minimal, in that neither has any proper subsets with the same properties. Singular therefore shares something with dual, just as dual shares something with plural, as illustrated in Table 1. Under this cross-classificatory organization, the shared features responsible for those natural classes are representationally on a par, equally and separately accessible for the grammar to reference in spellout. Either singular and dual (AAB) or dual and plural (ABB) may be neutralized to a syncretic form (realized with the same affix). However, singular and plural have no cross-cutting feature in common; this means singular-plural (ABA) syncretisms are excluded in cross-classifying feature systems.

Table 1: Schematic cross-classifying number feature system.

	singular	*nonsingular*
minimal	sg	du
		pl

This general organizational scheme for number features has been implemented in a wide variety of ways, as Table 2 demonstrates. Major proposals differ on feature valence, as well as on the nature of the relationship between features and syntactic terminal nodes.

Table 2: Some specific cross-classifying number feature proposals.

Privative features in geometries (Harley & Ritter 2002)

		Group
Minimal	sg INDV | Min	du INDV Min Group
		pl INDV | Group

Binary features in bundles (Harbour 2014)

	[+atomic]	[−atomic]
[+minimal]	sg [+a, +m]	du [−a, +m]
[−minimal]		pl [−a, −m]

Proposals may differ on feature valence: the shared pieces are sometimes monovalent features, mapping directly to the informal descriptions given above and in Table 1. For example, Harley & Ritter (2002) encode the nonsingularity of dual and plural as a privative Group feature, and the minimality of singular and dual as a privative Minimal feature. In other schematically identical theories, the shared pieces are valued binary features in the Jakobsonian tradition instead (Jakobson 1958; Hale 1973; Silverstein 1986). For example, Noyer (1992) and Harbour (2014) contrast the nonsingularity of dual and plural with the atomic singular using negative and positive values of [±atomic], and juxtapose singular and dual's minimality with the plural's containment of plural proper subsets using [±minimal]. Proposals may also differ on the morphosyntactic architecture these varying primitives are embedded in. Harley & Ritter's (2002) Group and Minimal features are dependent nodes in a feature geometry that encodes pronominal number, whereas values of Harbour's (2007; 2008; 2011b; 2011a; 2013; 2014; 2016; 2020[1]) binary features can form an unstructured bundle, occupying a single syntactic terminal.

None of these differences are crucial for the derivation of natural classes, or, therefore, for this paper. What is crucial is that all these theories make the intermediate status of the dual manifest through cross-classification; in deriving the dual by combining aspects of the singular and the plural that are on a par, these theories ensure that singular-dual (AAB) and dual-plural (ABB) are natural classes. For concreteness, in this paper I adopt the representations laid out in Harbour's body of work to exemplify cross-classifying approaches to singular, dual, and plural number (2007; 2008; 2011a; 2013; 2014; 2016; 2020). Further details of this representative theory are given in Section 1.4.1.

1.2 *Features in a linear hierarchy*

In contrast, theories that construct *linear hierarchies* of number features take dual to depend on, or contain, plural. Plural, in turn, depends on singular. These theorized featural dependencies are largely based on morphological markedness diagnostics (Greenberg 1966; Croft 1990; Noyer 1992; Nevins 2011; cf. Harbour 2011a; also see discussions of semantic markedness in Sauerland 2008; Bale et al. 2011; Bobaljik et al. 2011; i.a.). Typologically, as is well known from Greenberg's (1963) inventory-based implicational universal, the dual category only appears in languages that also have plural marking (74, Universal 34). Formally, the overt coding of dual marking is often more complex than that of the plural (Nevins 2011: 418; Smith et al. 2019; also see (7)), and the plural is always overtly marked in some way, as compared to the often zero-marked singular (Corbett 2000). The more complex dual requires plural as its foundation, just as plural requires the least marked singular as its foundation, as illustrated schematically in Figure 1. Under this linear dependency organization, any reference to a more complex number value definitionally also catches the next value down. Either the immediately contiguous dual and plural (ABB) or the immediately contiguous plural and singular (ABA) values may participate in a syncretism. However, singular and dual are noncontiguous; this means singular-dual (AAB) syncretisms are excluded in linear dependency feature systems.

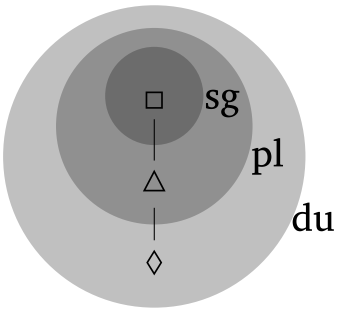
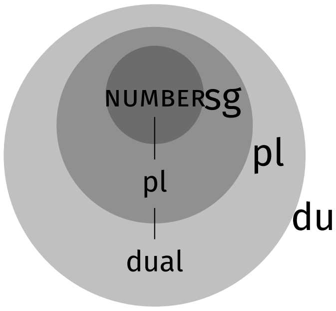
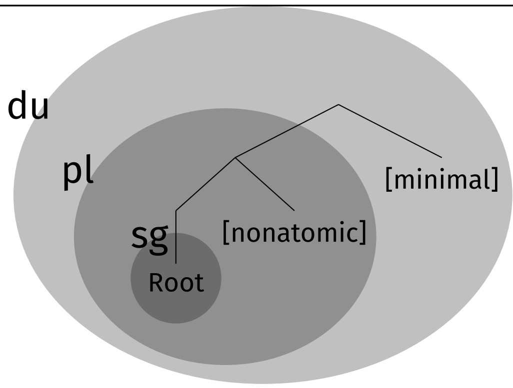


Figure 1: Schematic linear hierarchy number feature system.

[1] Harbour (2020) reaffirms that [±atomic] and [±minimal] syntactically cooccur on the same head in any language with a full three-way singular, dual, and plural distinction in nominal morphemes. Only for the specific phenomenon of Frankenduals — where a language that only has a two-way morphological distinction between singular and plural, nevertheless constructs a composed dual by combining plural (i.e. [−atomic]) nominal marking and singular (i.e. [+minimal]) verbal (or other more peripheral) marking — does Harbour (2020) merge [±atomic] and [±minimal] features in separate syntactic locations. The existence of this exceptional pattern serves as a reminder that it is not the *syntax* of these features in any given language or account that is crucial for this paper. Rather, it is their cross-classificatory organization of the number values: that they fundamentally entail that "dual must lie at the featural intersection of two natural classes, one with singular, the other with plural" (Harbour 2020: 84).

As with cross-classification, there are diverse implementations of a linearly hierarchical organizational scheme for number features, as Table 3 shows. These proposals also differ on feature valence and the syntactic arrangement of feature representations.

Table 3: Some specific linearly hierarchical number feature proposals.

Privative features in geometries (Harley 1994)	Privative features in containment (Smith et al. 2019)
	

For example, Harley (1994) proposes privative representations structured in a geometry, but with a different set of dependency relationships than in Harley & Ritter (2002). Unlike that later proposal, the monovalent component pieces identifying singular and plural are not equal dependents of a separate node, and dual is not formed from their combination; pl is instead asymmetrically dependent on the default singular number, and dual is in turn asymmetrically dependent on pl. The linear hierarchy of Figure 1 can be encoded across separate syntactic heads rather than in a geometrically structured bundle, too. For example, Smith et al. (2019) separate each number feature onto its own distinct terminal node. Smith et al. (2019) additionally suggest that the binarity of Harbour's features may be limited to syntactico-*semantic* interpretation; if only the marked value of each feature is *morphosyntactically* visible, then attaching the marked [nonatomic] feature to singular will create a larger, more marked plural; attaching the marked [minimal] feature to that tree in turn will create a larger, more marked dual. Dual, then plural, then singular, structurally contain one another in an ordered hierarchy of syntactic trees.

Again, none of these architectural differences matter for which natural classes are derived. What is crucial is that all these theories encode the asymmetric markedness of dual, then plural, then singular as linear dependency. In embedding less marked numbers directly within more marked numbers, these theories ensure that the contiguous dual-plural (ABB) and singular-plural (ABA) are natural classes. For concreteness, in this paper I adopt the morphosyntactic representations put forth in Smith et al. (2019: Section 4.3.3) for singular, dual, and plural number. Further details of this example theory are given in Section 1.4.2.

1.3 *Number syncretism in artificial languages*

This paper presents an artificial language learning experiment that harnesses this connection between feature organization and representable syncretisms to probe participants' mental representations of number features. Cross-classifying features cannot directly represent singular-plural (ABA) syncretism, whereas linearly organized features cannot directly represent singular-dual (AAB) syncretism. In order to probe which of these possible mental representations of number are actually in use in learning, 149 adults were trained on one of three fragment grammars, to investigate their relative learnability. All three grammars had a three-way number distinction between singular, dual, and plural on nouns, but each displayed a different syncretism in verbal number agreement (dual-plural ABB, singular-dual AAB, or singular-plural ABA). Participants' relative success in learning each syncretism was measured by their performance on a referent selection task based only on the verbal number agreement marker. Relative accuracy across the three grammars reveals which pairs participants treat as natural classes and which they consider featurally random and therefore harder to learn. In this way, the experiment weighs opposing notions of morphological feature organization.

I now present one such miniature language. These are artificial languages spoken by alien characters (described in more detail in Section 2.2.1), but real languages with this kind of syncretism do exist (discussed in Section 4.3).

Each artificial language had the same three-way distinction between singular, dual, and plural number on nouns referring to fruits or vegetables, as in example (1). Using lexical nouns rather than pronouns avoids complicating interactions with person features, which can change the meanings of certain number features (Noyer 1992; Harbour 2016) and are themselves in need of study in artificial language contexts (Maldonado & Culbertson 2020). Restricting noun meanings to a single semantic area focuses learner attention on the morphosyntactic category of interest: number.

(1) Three-way number contrast on lexical nouns

a. deet-cha
banana-SG
'(a) banana'

b. deet-po
banana-DU
'(2) bananas'

c. deet-fi
banana-PL
'(>2) bananas'

The three values in this system allow for three groupings, one of which is dual and plural (to the exclusion of singular), or ABB. Example (2) demonstrates the grammar making use of that grouping in a conditioned syncretism. The invented verb's suffix agrees in number with its object. However, this verbal context neutralizes the elsewhere well-established distinction between dual and plural. The neutralized form /-ku/ appears in both contexts in (2b)-(2c). This is ABB dual-plural syncretism, which the following subsection will walk through two competing analyses of.

(2) ABB number syncretism on imperative verbs

a. Bice-**te** deet-**cha**!
beam.up-SG.O banana-SG
'Beam up the banana!'

b. Bice-**ku** deet-**po**!
beam.up-DU.PL.O banana-DU
'Beam up the (two) bananas!'

c. Bice-**ku** deet-**fi**!
beam.up-DU.PL.O banana-PL
'Beam up the (more than two) bananas!'

It is critical that this neutralization in one domain comes with the full differentiation of all three numbers in another. This establishes that the features necessary for the three-way distinction in (1) are still in play for the verbal agreement in (2). They cannot just conflate the meaning of dual with the meaning of plural everywhere, potentially using a simplified set of feature structures. Such a conflation is found in English: the potential impact of participants' experience with the dual-plural conflation in English on their performance with the ABB conditioned syncretism is discussed in Section 4.4.1.

1.4 *Example implementations of cross-classification and linear dependency*

The remainder of this section provides concrete implementations of the opposing approaches to feature organization being compared. As Sections 1.1 and 1.2 have already emphasized, no single aspect of the morphosyntactic architecture — whether that is morphosyntactic feature valence, the syntactic organization of features into terminals, or the mechanisms for spellout — is solely responsible for the overall cross-classificatory or linearly dependent character of a given proposal for singular, dual, and plural number. Nevertheless, certain combinations of architectural choices are particularly compatible with one or the other kind of relationship among morphological values, and sometimes have been expressly designed to be that way. The detailed discussion of a specific, fully worked-out proposal within each camp therefore serves to illustrate the role that substantive questions about natural classes and feature organization have had in the development of different theoretical frameworks (Embick 2016).

1.4.1 Harbour (2014; 2016) and Distributed Morphology: an example of cross-classification

Distributed Morphology (DM) is a realizational framework that allows for cross-classificatory features to derive morphological natural classes (Halle & Marantz 1993). As a Late Insertion, Separationist theory, DM provides phonological form, like the three distinct number exponents in (1), to abstract morphemes in a syntactic structure. Whatever the featural content of an abstract morphological terminal node, any exponent linked to a subset of those features may compete for insertion, with the most specific competitor winning.

No defining aspect of the canonical DM architecture makes stringent demands on how those abstract morphemes are constructed, or which kinds of features they are constructed of. Multiple features can occupy a single syntactic terminal as a bundle, or a lone feature can. Some features may be valued binary features (either positive or negative) and some may be privative. As a result, Noyer (1992) and Harbour's (2014; 2016) proposal of two binary number features[2] in a bundle can be comfortably couched within DM. As the rightmost cell of Table 2 illustrates, [±atomic] definitionally separates out (singular) atoms, and [±minimal] picks out sets that have no proper subsets with the same properties. These two available binary features generate 2^2, or four feature bundles that may be terminals. Only three of these four combinations are semantically viable, generating singular, dual, and plural.[3] Each of these combinations of the available binary features receives a separate exponent in the nominal number suffixes in (1), demonstrating that insertion rules for Vocabulary Items may reference all of the features found in a terminal's feature complex, as in (3)a-c. For example, the Vocabulary Item in (3b) would exactly match the terminal bearing both [−atomic] and [+minimal] features in (4b), and thereby expone the nominal dual number marking seen in (1b) and (2b), /-po/.

[2] Noyer's (1992) initial decomposition of singular, dual, and plural into [±singular] and [±augmented] directly inspired Harbour's account: in earlier work, Harbour (2007; 2008; 2011a) uses Noyer's (1992) terminology. Harbour (2013; 2014; 2016) rechristens [±singular] as [±atomic], and names the inverse of [±augmented] to be [±minimal]. In all critical respects for this paper, these accounts are equivalent: they pick out the same natural classes, and both use underspecification to spell out these natural classes as syncretic.

[3] The semantics of the three viable feature bundles are as follows: singular atoms are definitionally picked out by [+atomic], and trivially satisfy [+minimal] because they have no proper subsets of any kind. They cannot combine with [−minimal]. The [−atomic] non-singulars are split into the [+minimal] dyads, which have no plural proper subsets, and the [−minimal] pluralities, which do.

(3) Cross-classifying features (Harbour): Vocabulary Items for ABB grammar

Nominal number suffixes:

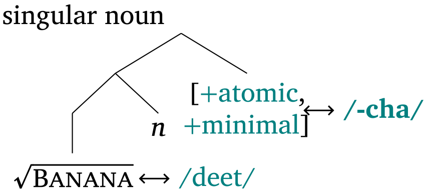
a. singular:
   [+atomic, +minimal] ⟷ /-cha/
b. dual:
   [–atomic, +minimal] ⟷ /-po/
c. plural:
   [–atomic, –minimal] ⟷ /-fi/

Verbal number agreement suffixes:

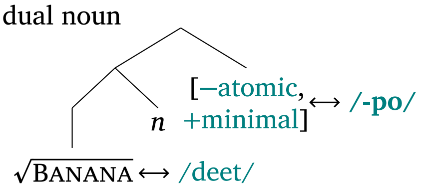
d. singular:
   [+atomic, +minimal] /v ⟷ /-te/
e. dual/plural:
   [–atomic] /v ⟷ /-ku/

(4) Cross-classifying features (Harbour): ABB derivations

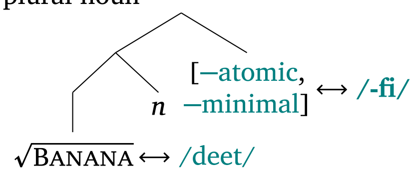
a. singular noun



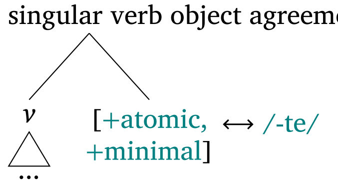
b. dual noun



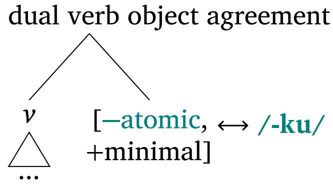
c. plural noun



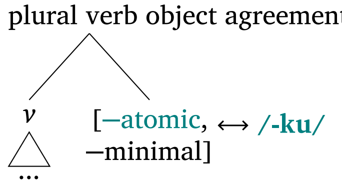
d. singular verb object agreement



e. dual verb object agreement



f. plural verb object agreement



Distributed Morphology's spellout mechanisms translate Harbour's proposal about the compositional semantics and bundled syntax of number features into the morphological natural class predictions that this paper is concerned with. Though VI insertion rules may reference all of the features in a bundle, they do not have to; Distributed Morphology's Subset Principle allows Vocabulary Items to be featurally underspecified with respect to syntactic terminals. If multiple VIs are able to match a subset of the features in a syntactic terminal, the competition is resolved through specificity: the VI with the greatest number of matching features will be inserted. In this way, featural underspecification of Vocabulary Items can capture the verbal number agreement syncretism in (2). In that verbal context, a single underspecified Vocabulary Item referencing only [–atomic] (3)e can be inserted for both dual (4e) and plural (4f). No more specific Vocabulary Item exists to compete for insertion in that verbal context. The singular feature complex is matched by a separate [+atomic] Vocabulary Item; whether or not it contains [+minimal], (3)d is the only VI that matches any features of (4d). Subset-based spellout and underspecification of individual Vocabulary Items lets the grammar pick out the dual-plural natural class to be realized as a shared-feature syncretism (Kramer 2016: 97).

Similarly, this system of Noyer-Harbour features within DM can also realize a singular-dual (AAB) shared-feature syncretism. A Vocabulary Item referencing only the [+minimal] feature in the verbal context (5)d would pick out both the singular (4d) and dual (4e) numbers. Note that the syntactic representations in (4)d-f remain constant between both the ABB and AAB grammars; both the [±atomic] and [±minimal] features must apply to derive the correct semantics for the three numbers, even if only one is referenced in a particular spellout context.

(5) Cross-classifying features (Harbour): Vocabulary Items for AAB grammar

Nominal number suffixes:

a. singular:
   [+atomic, +minimal] ⟷ /-cha/
b. dual:
   [–atomic, +minimal] ⟷ /-po/
c. plural:
   [–atomic, –minimal] ⟷ /-fi/

Verbal number agreement suffixes:

d. singular/dual:
   [+minimal] /v ⟷ /-te/
e. plural:
   [–atomic, –minimal] /v ⟷ /-ku/

Conversely, DM cannot represent singular-plural (*ABA) neutralization as a shared-feature syncretism using any value of Harbour's features. No underspecified Vocabulary Item can target singular and plural with the same exponent; no value of either feature is shared between the two feature complexes (though see Section 4.2.2 for further discussion on underspecification).

In sum, placing Harbour's cross-classifying features within DM derives ABB and AAB patterns of shared-feature syncretism, but not *ABA via underspecification as just described.

1.4.2 Smith et al. (2019) and Nanosyntax: an example of a linear hierarchy

Given Distributed Morphology's theoretical permissiveness on syntactic terminal creation, its subset-based spellout principles could straightforwardly apply to a linear-dependency-based representation of number (e.g. Harley 1994, discussed in Section 1.2) and accordingly derive ABB and ABA syncretisms instead. That is, whether singular-plural (ABA) turns out to be less learnable than other syncretisms, as cross-classification predicts, or singular-dual (AAB) does, as linear dependency predicts, the key architectural tenets of DM are not at stake. This flexibility has made DM a big-tent framework compatible with various kinds of featural representations, which could be different across morphological categories like number, gender, person, and case.

In contrast, one emerging view of the architecture axiomatically prescribes linear dependency relationships among morphological values across all categories: Nanosyntax (Starke 2010). In Nanosyntax, each feature is its own syntactic terminal; morphological values built from those features correspond to subtrees in structural containment relations. Such syntactic subtrees, not just terminal nodes, can be matched by any lexical superset tree that contains it, with the most specific match winning. The nature of syntactic terminals in Nanosyntax thus directly encodes linear hierarchy "as a core hypothesis about the architecture of grammar," and its superset-driven spellout principles derive adjacency-based morphological natural classes (Caha 2019: 3). Unlike DM, then, Nanosyntax takes a particular feature organizational scheme as essential to its architecture. Linearly hierarchical number features are only compatible with ABB and ABA number syncretisms; if singular-dual (AAB) syncretism turns out to as learnable as those others, the Nanosyntactic framework itself can be called into question.

In the hopes that it will be an instructive test of the framework, then, this paper casts Smith et al.'s (2019) analysis of singular-dual-plural number systems within Nanosyntax. Although Smith et al. themselves allude to using DM as a framework (2019: 1034), they ultimately propose representations that require adding parametric assumptions to DM, but already conform directly to core requirements of Nanosyntax (2019: 1074–5). Using Nanosyntactic principles to spell out those representations just converts linear adjacency into syncretism of a natural class more directly.

In Nanosyntax, features have no values: they are either present or absent, as their own syntactic terminals. Although Smith et al. directly adopt Harbour's semantic definitions of [±atomic] and [±minimal] to capture the meanings of the three numbers, they convert them into very different kinds of morphosyntactic objects, which conform to that Nanosyntactic dictate.

First, rather than having both positive and negative values of each feature available in the morphology to determine spellout, only one value of each feature is actually represented, as in (6) (Smith et al. 2019: 1074–5: (37)).[4] The feature value that is represented is the one that derives the more marked number. This representation transparently captures patterns where the dual contains the plural, as in Manam (see (7)). As dual is cross-linguistically rarer than the singular-plural split, it follows that both of the feature values that compose the dual — [−atomic] and [+minimal] — must be marked, and therefore are represented in the morphosyntax. The two exponents /-di/ and /-ru/ in the Manam dual demonstrative (7c) transparently realize those heads. Conversely, the unmarked [+atomic] and [−minimal] are absent from the representation and therefore unavailable for exponence. The plural number only contains [−atomic] (and not the unmarked [−minimal]), and so surfaces only with /-di/ in (7b), and the least marked singular number has no featural content (or number exponence) at all in (7a).

(6) sg: ~~[+atomic]~~ pl: [−atomic] ~~[−minimal]~~ du: [−atomic] [+minimal]

(7) Manam demonstratives show “plural” contained by “dual”

a. áine ŋára	b. áine ŋára-di	c. áine ŋára-dí-a-ru
woman that	woman that-[−a]	woman that-[−a]-L-[+m]
‘that woman’	‘those women’	‘those (two) women’

Second, Smith et al. (2019) separate the marked [−atomic] and [+minimal] features out of Harbour’s bundles onto distinct terminal nodes. Those terminals are ordered in a universal functional sequence (f-seq). This, along with the absence of the nonmarked value of each feature, means that less marked number values are literally contained inside other morphosyntactically larger, more marked number values. Smith et al. (2019) argue that that is the relationship between singular, plural, and dual — in that order.[5] Both morpheme order in languages like Manam, and the semantic requirement that [±atomic] must apply before [±minimal] (Harbour 2016: 206) lead them to propose that [+minimal] should project above/after [−atomic].

Note that Smith et al. (2019) themselves do not commit to the Nanosyntactic claim that the containment structures in (6) are universal for singular-dual-plural languages. They allow for the possibility of parametric variation in the morphological markedness of [±minimal], resulting in some languages where plural contains dual (instead of the other way around) (2019: 1072–4). Languages that flip the containment order in this way would also flip predictions about possible syncretisms. Since admitting parametric variation in this way would erase any strong cross-linguistic predictions about possible syncretisms, the remainder of this paper only considers the baseline linear feature containment order of du > pl > sg discussed by Smith et al. (2019), given in (6).

Given these choices, the number values of singular, dual, and plural are represented by morphosyntactic subtrees containing privative features on separate heads, as in (8).[6]

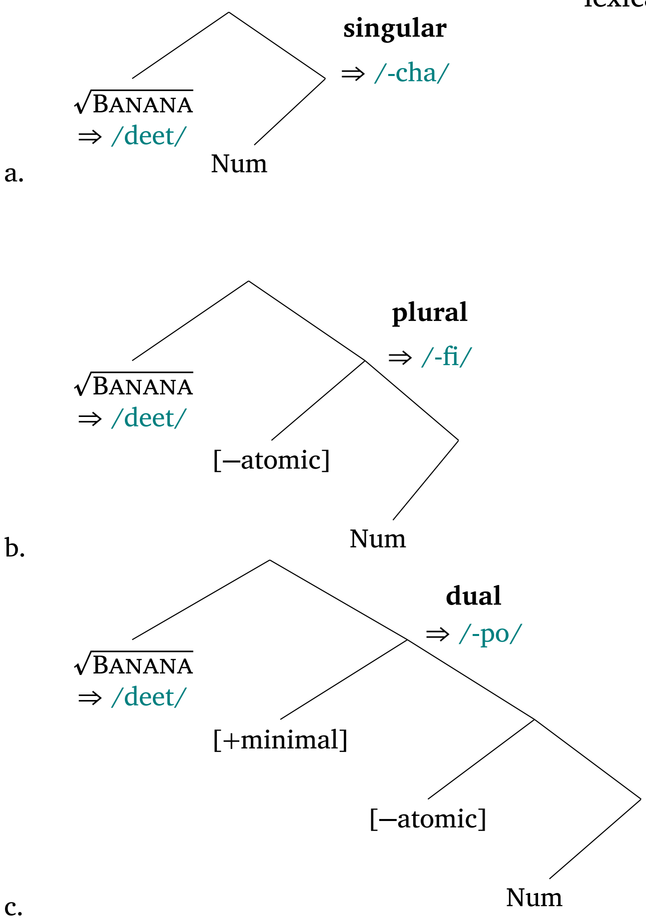
[4] Note that representations using the Smith et al. (2019)-style features in examples (6)-(13) retain a plus on [+minimal] and a minus on [−atomic]. This reflects the semantic interpretation of these features, as well as Smith et al.’s own notation and diagrams (Smith et al. 2019: 1074–5: (37)). However, N.B. that the opposite value of each feature is not actually available to be syntactically represented; as far as spellout is concerned, the features are functionally privative.

[5] I translate Smith et al.’s (2019) singular-plural-dual (ACB) order of containment to the singular-dual-plural (ABC) order already introduced, whenever discussing syncretisms.

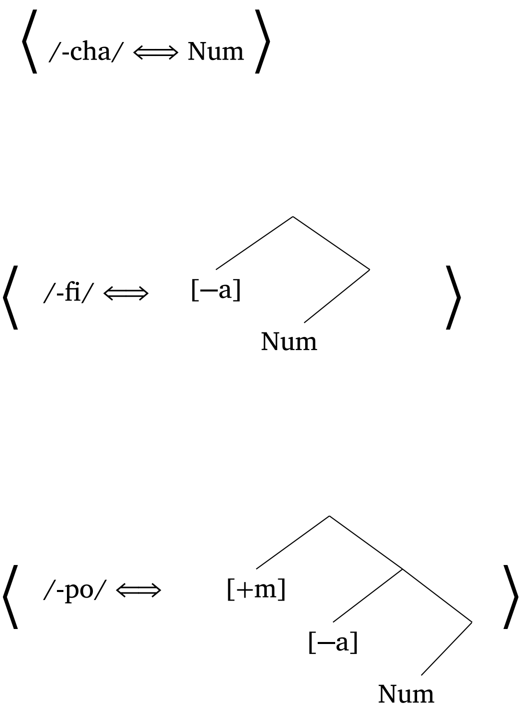
[6] The roots in the syntactic derivations given in (8), (10), and (12) have been moved from their base-generated, innermost positions to specifier positions at the top of the tree. These Nanosyntax-style spellout-driven movements leave no traces, and happen in order to “generate new constituents that are candidates for lexicalization” (Baunaz et al. 2018: 37, Starke 2018). In these cases, the movement of the root leaves behind the number sequence, which is then allowed to be spelled out as a suffix.

Although I use [+] and [−] symbols in the notation for the [−atomic] and [+minimal] heads, following Smith et al.'s own notation and diagrams (2019: 1074–5: (37)) and reflecting how they are semantically interpreted, only [−atomic] and [+minimal] are represented in the morphosyntax and are visible for spellout. This makes them morphosyntactically privative. Each of these subtrees receives a separate pronunciation as the nominal number suffixes in (1)-(2), demonstrating that exponents must be able to be inserted at non-terminal syntactic nodes. Morphophonological pronunciations are stored with subtrees of their own, as in (9), and Spellout is the process of matching the syntactic tree to those lexical entry subtrees, working cyclically outwards from the Root. A given lexical subtree may match all the nodes of the morphosyntactic tree up to that point, in which case it will be inserted, as the three lexical entry trees in (9) are.[7]

(8) Linearly dependent features (Smith et al. 2019): nominal derivations



(9) Linearly dependent features (Smith et al. 2019): lexical entries for noun affixes



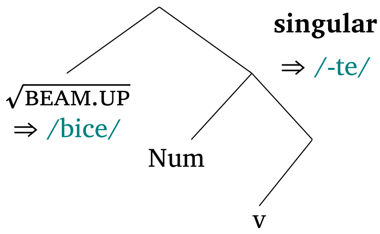
[7] Note the contrast between the structures in (8)-(9) and the structures in the rightmost cell of Table 3. In (8a) and (9), the nominal singular is represented by a Num featural head, whereas it is represented by a 'Root' head in Table 3. Across all representations, the plural and dual are then built atop that singular structure.

If singular were equivalent to the Root, then the Superset Principle would require the plural and dual structures which contain it to have pronunciations that either (a) also include the singular form overtly within them (this is equivalent to there being a single stem in all numbers, with a zero singular exponent), or (b) overwrite the singular form with a different suppletive root(s) altogether (like English 'I' and 'we'). Crucially, if there is no additional non-Root head within the singular, then there is no way to derive a single stem (like /deet/ for the nominal root √BANANA) lexicalized separately from three separate number affixes. This situation is evidently not an issue in Smith et al. (2019), as they are precisely trying to characterize the triggers of (pronoun) root suppletion: for them, the "Root" in the rightmost cell of Table 3 refers to "the person formative" and is "intended loosely [to refer to] the most deeply embedded morpheme in the pronoun and the one that undergoes suppletion" (2019: fn9). But for non-suppletive lexical nouns or agreement series with overt affixes in all three numbers, as in the artificial languages used in this experiment, this structure is insufficient. Accordingly, I adopt the Num head as the representation of singular, forming the innermost base of all the number values, following representations given in Starke (2020).

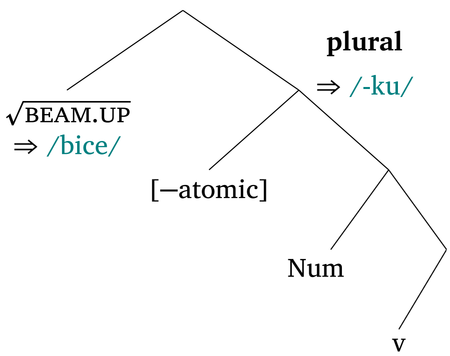
Though lexical subtrees may match morphosyntactic subtrees exactly, they do not have to; Nanosyntax's Superset Principle allows lexical entry subtrees to contain more material than the syntactic structures that they spell out. As long as the morphosyntactic tree is contained within the lexical entry subtree, matching occurs and insertion is possible (Starke 2010; Baunaz et al. 2018; Caha 2019). If there are multiple lexical entries that contain a syntactic structure, the competition is resolved through specificity and the elsewhere principle (also called Minimize Junk): "if several lexical items match the root node, the candidate with least unused nodes wins" (Starke 2010: 4). This system can capture the ABB verbal number agreement syncretism in (2). By the Superset Principle, a lexical subtree that contains the syntactic representations of the structurally largest number value — the dual — can be inserted for all the number values it contains. Such a lexical subtree is given in (11) with the morphophonological form /-ku/. If it were the only lexical entry available, it could be inserted for any of the syntactic trees in (10), as it lexicalizes the unbroken line of heads forming singular, plural, and dual (and the verbalizing head[8]). However, there is also a lexical subtree with the phonological material /-te/ that matches the singular syntactic structure in (10a) exactly; it wins the competition for singular. This derives a dual-plural ABB syncretism.

(10) Linearly dependent features: ABB verbal derivations

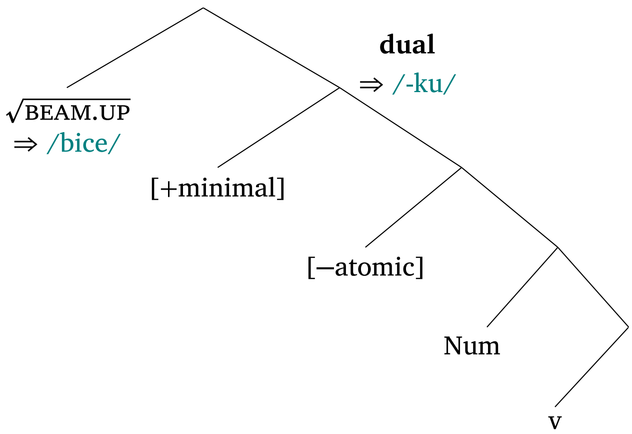
a.



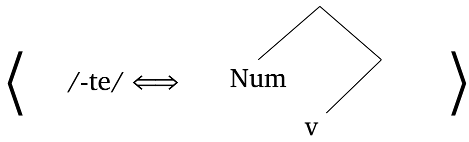
b.

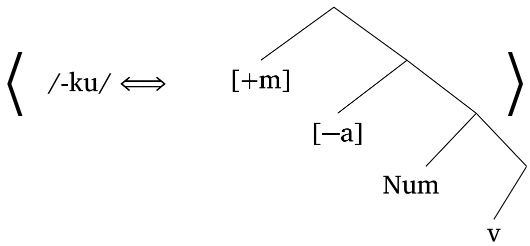


c.



(11) Linearly dependent features: lexical entries for ABB verbal affixes



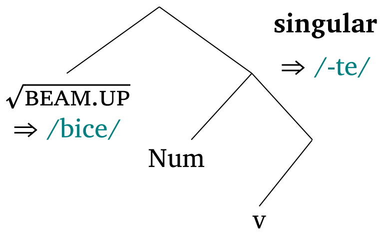


[8] This is a stand-in for whatever lower head actually defines verbal agreement, and thereby differentiates number exponence in a verbal agreement complex from number exponence in its base-generated position on nouns (pc, Hagen Blix).

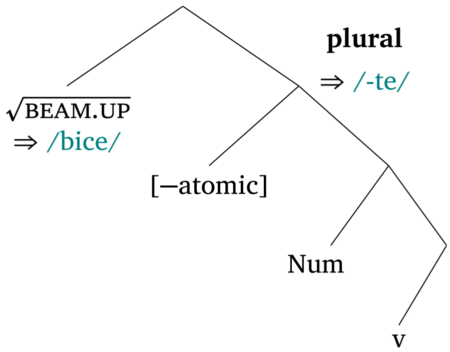
Nanosyntax can map Smith et al.'s (2019) containment hierarchy to ABA syncretism in a similar way: the lexical subtree containing the structure up to [−atomic] and pronounced /-te/ in (13) can lexicalize both singular and plural, winning over the superset /-ku/ subtree because it contains fewer unused nodes in both cases. For dual, the /-ku/ subtree matches exactly and is inserted.

(12) Linearly dependent features: ABA verbal derivations

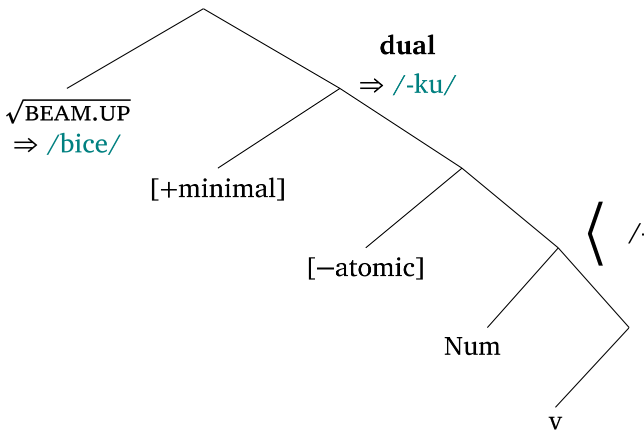
a.



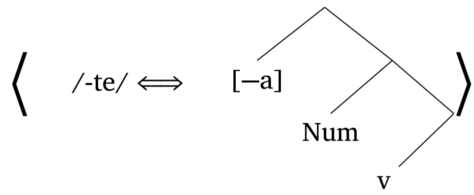
b.

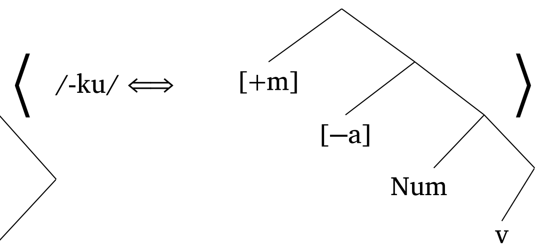


c.



(13) Linearly dependent features: lexical entries for ABA verbal affixes





On the other hand, Nanosyntax has no way to derive *AAB using Smith et al.'s (2019) containment hierarchy. To do so, you would need a subtree to match only plural (B) and one to match both singular and dual (A). However, a subtree only containing [−atomic] would not be able to match the plural syntactic subtree, and the maximal subtree (like /-ku/ in (11) and (13)) meant to target singular and dual would also end up spelling out plural. Only contiguously adjacent number values can be spelled out together for suppletion or syncretism, and "nonadjacent syncretisms" — in this case, singular and dual — "are excluded in principle...by the *ABA theorem" (Baunaz et al. 2018: x). Note that the *ABA theorem discussed in Smith et al. (2019) and Baunaz et al. (2018) corresponds to the singular-dual AAB syncretism in this paper; this is because Smith et al. refer to patterns of suppletion in the order singular-plural-dual (ACB) to match the structural containment they propose in (6). I have converted these patterns into the singular-dual-plural (ABC) order used throughout this paper.

In sum, placing Smith et al.'s (2019) containment structures for singular, plural, and dual within Nanosyntax predicts ABB and ABA patterns of syncretism, but not *AAB.

1.5 *Theoretical predictions for syncretism patterns*

The theories of number features just surveyed agree that the dual and plural values form a morphosyntactically relevant natural class. The ABB syncretism introduced in the artificial language in Section 1.3 is predicted to be learnable. However, the theories make opposing predictions for the natural class membership of the singular, with cross-classifying features predicting a disadvantage for learners exposed to the singular-plural (ABA) syncretism grammar, whereas linearly hierarchical features predicts one for learners of the singular-dual (AAB) syncretism. These predictions are summarized in Table 4.

Table 4: Natural class predictions by theory.

		ABB du-pl	AAB sg-du	ABA sg-pl
Cross-classifying	[±a, ±m]	✓	✓	✗
Linearly hierarchical	[[[]–a]+m]	✓	✗	✓

2 Experiment

2.1 *Design*

To test whether the natural classes predicted by a given number feature theory are innately active in learning, this experiment asks participants to learn a syncretic pattern that groups two number values. In this between-subjects ease-of-learning experiment (Culbertson et al. 2017, Maldonado & Culbertson 2020: 8), participants were randomly assigned to one of three grammar conditions: ABB, AAB, or ABA, illustrated in (14). As discussed above in Section 1.3, all grammar conditions have lexical nouns that bear singular, dual, and plural affixes, to impart the three-way number distinction that ensures all relevant representations are morphosyntactically active. Each grammar condition contextually neutralizes that three-way distinction on verbs' object agreement suffixes, differing on which two forms are syncretic: dual-plural (ABB), singular-dual (AAB), or singular-plural (ABA). Participants were trained on one of these number systems, then tested on their comprehension of the nominal and verbal number affixes through forced choice referent selection tasks. Relative participant accuracy across the grammar conditions provides a measure of the learnability of each syncretism, and the activity of feature-based natural classes in learning artificial number systems.

(14) **Experimental conditions (grammar)**

a. **ABB**

nouns	verbs
sg	A
du	B
pl	B

b. **AAB**

nouns	verbs
sg	A
du	A
pl	B

c. **ABA**

nouns	verbs
sg	A
du	B
pl	A

2.2 *Materials*

The experimental materials included audio from three versions of a miniature artificial language. Their referents were represented by visual animations and scenes, relating to two aliens with a spaceship interacting with different quantities of fruit.

2.2.1 Language

As already demonstrated in Section 1.3, the artificial language is a suffixing language that marks a three-way distinction in the nominal domain and partially neutralizes it in verbal agreement. All three versions of the language contain 6 nominal stems referring to different kinds of fruit, given in Table 5, and 2 verbal stems referring to the actions ‘beam up’ or ‘throw away’ (performed by the aliens’ spaceship), given in Table 6. The forms for these lexical items are 8 CVC forms from Vitevitch & Luce’s list of English-like monosyllabic nonwords with high phonotactic probability and high neighborhood density (1998; 1999: Appendix A, following Pylkkänen et al. 2002).

Table 5: Nominal stems.

pem [pem]	carrot
seeg [sig]	pear
dite [daɪt]	lemon
lun [lʌn]	cherry
deet [dit]	banana
feek [fik]	apple

Table 6: Verbal stems.

pim [pɪm]	beam up
bice [baɪs]	throw away

Table 7: Nominal number suffixes.

-cha [ʧa]	singular
-po [poʊ]	dual
-fi [fi]	plural

Table 8: Verbal number agreement suffixes.

<table>
<tr><th></th><th>ABB</th><th>AAB</th><th>ABA</th></tr>
<tr><td>sg</td><td>A: -te [teɪ]</td><td rowspan="2">A: -te</td><td>A: -te</td></tr>
<tr><td>du</td><td rowspan="2">B: -ku [ku]</td><td>B: -ku</td></tr>
<tr><td>pl</td><td>B: -ku</td><td>A: -te</td></tr>
</table>

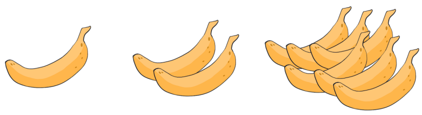
Nouns bear one of three suffixes to mark singular, dual, and plural, as in Table 7, and verbs track the number of their object via one of two suffixes (represented as A and B in Table 8). The meaning of those verbal suffixes depends on the grammar. The forms of these number suffixes are 5 CV syllables constructed from 5 different common English onsets and 5 different common English vowels to ensure clear distinctions (see Culbertson et al. 2019: fn8). Care was taken in the assignment of these CV forms to the 5 number suffix meanings (given the range of coda consonants found in nominal vs. verbal stems) to avoid geminate clusters.

Utterances in the language were all presented aurally in the training and test phases of the experiment, but Tables 5 through 8 also include the orthographic transcriptions used in scripts for the voice actors (and in nonce-word discrimination items used as attention checks in the experiment, described in Section 2.3 below). The auditory stimuli were recorded by three American English speakers, who were instructed to speak using American English pronunciation and phonotactics, including initial stress. All recordings were made in a soundproof booth using a Zoom H4n Handy Recorder and digitized, edited to a similar length as other items, and converted into mono in Praat acoustical analysis software (Boersma & Weenink 2021). Additionally, two test audio stimuli were constructed by splicing the disambiguating, number-marked nominal out of imperative sentences and replacing them with a 1.0-second-long pulse train waveform synthesized in Praat. This left a verb bearing a potentially syncretic number agreement marker followed by obscuring noise in the usual position of its nominal object. The resulting clips were saved in .mp3 for compatibility with a wide range of web browsers.

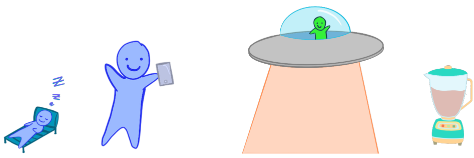
2.2.2 Visuals

Referents for the singular, dual, and plural numbers were color drawings of eight different kinds of produce (apples, bananas, carrots, cherries, lemons, watermelons, oranges, and pears) created in Adobe Draw, shown in (15). Each lone fruit represented a singular form. A second copy of the fruit drawing layered into the same image represented a dual form. To represent plural forms, a random number between 6 and 9 was generated for each type of fruit. This higher numerical range allows us to avoid consistently matching any attested exact number values grammaticalized in the world's languages (i.e. the trial (Corbett 2000: 21–22)), or specific cardinalities that participants might unduly focus on. Moreover, the range is well above the adult subitizing range, outside of which "fast, effortless, and accurate...enumeration of a visual set of items...becomes slower, effortful, and error-prone" and which Nevins & Marty have suggested could form the cognitive basis for the paucal number value (Under revision: 19).

(15) Visual stimuli: fruit



(16) Laru (speaker) & Nika (addressee)



Animations showing two alien characters and their accoutrements (shown in (16)) interacting with these quantities of fruit expressed the verbal meanings of the language. The blue alien, named Laru, presented all the artificial language utterances in the training phases, and provided an anthropomorphic focal point throughout the experiment. Their movements were also animated to provide feedback as participants learned the language through practice referent selection items. Laru was equipped with a computer tablet, which both provided artificial language stimuli in the test phase and added the obscuring buzzing noise that isolated the syncretic verb forms. The green alien, named Nika, served as Laru's addressee and piloted the spaceship to accomplish all the actions Laru asked of them. Additional images illustrating aspects of the frame story, such as Laru falling asleep and a blender for a fruit smoothie, were also created.

2.3 *Procedure*

Testing took place online on participants' keyboard- and speaker-equipped personal computers. Amazon Mechanical Turk presented a written consent statement, general instructions, and a link to the Pavlovia website, which administered the PsychoJS experiment (Peirce et al. 2019) as follows.

Participants were introduced to two friendly visiting aliens who love trying local fruits, via text and static images of each character. The green alien named Nika pilots a spaceship with the ability to beam things up or toss them away. The blue alien named Laru stands in the foreground of the screen; participants are instructed to try to learn the aliens' language from them.

Two basic sound checks preceded the four main phases of the experiment. To ensure they were listening to the audio and attending to the training and task, participants first heard a sound clip of an English sentence naming a number key to press on their keyboard (following D'Onofrio 2014). They then were presented with a sound clip of a single English word and asked to pick the word from four options displayed on the screen using their arrow keys. The four options were monosyllabic English words with the vowel nuclei found in the artificial language's morpheme inventory (i.e. [oʊ], [aɪ], [u], [i]). The four words were selected to be roughly as frequent as one another (between the 500th and 750th most frequent nouns in the Corpus of Contemporary American English (COCA, Davies 2009)). Each participant had to type the correct key *and* make the correct choice across the two sound check items for their data to be included in analysis.

The main experiment was divided into four phases: training and testing on nominal number, and training and testing on imperative verbs' number agreement markers.

2.3.1 Nominal number

Each training phase was divided into an exposure sub-block and a pick-the-right-one sub-block. Figure 2 represents the nominal training phase. Participants first heard nouns with the three number suffixes, with pictures of one, two, or many of the fruit accompanying the audio (with 12 repetitions). They were asked to repeat what they heard each time, with the option to advance to the next screen via the space key only becoming available 1.0 seconds after the audio clip finished playing. This exposure portion was followed by practice at the referent selection task: participants were asked to choose between two quantities of fruit based on audio, and given feedback on their choice (4 trials). The correct/audio number was randomly chosen from the three values, and the displayed numbers were taken at random from the possible pairs of {sg/du, sg/pl, du/pl}, with the correct referent image randomly placed on the left or right of the screen. Participants were asked to make the correct selection using their 'left' and 'right' arrow keys, which became active only after the audio had completed playing. They were told that Laru, the teaching alien character, would show them the right answer, even if they made the wrong choice. Feedback consisted of Laru moving over to the correct choice.

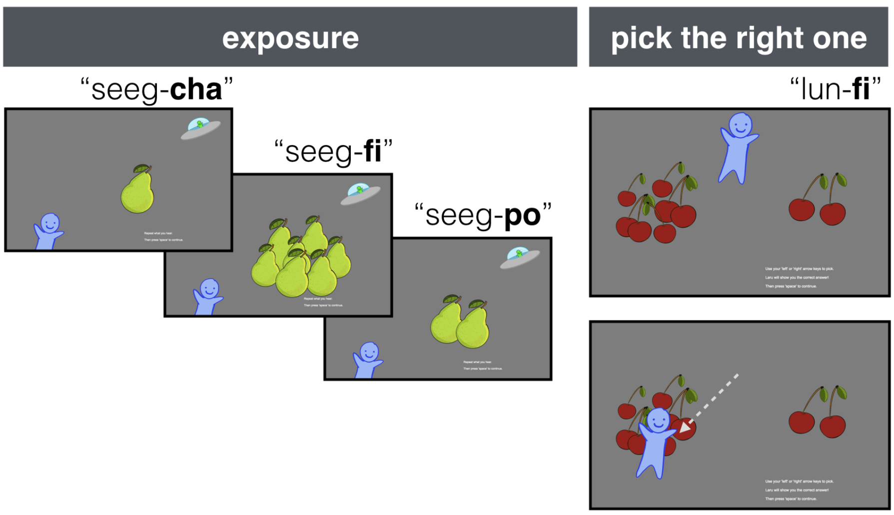


Figure 2: Nominal number training phase.

Each test phase measured success in learning the number morphology of the language via a forced-choice referent selection task, with no feedback. Figure 3 represents the nominal test phase. To validate participants' success in learning the three-way number distinction, they were asked to choose between three quantities of fruit based on audio of noun forms with number endings (8 trials). No feedback was given on their choices. Besides the three-way choice and the lack of feedback, these test items were otherwise completely parallel to the pick-the-right-one portion of the nominal training phase, with random selection of the correct number, random image placement on the left, middle, or right of the screen, and arrow key answer selection.

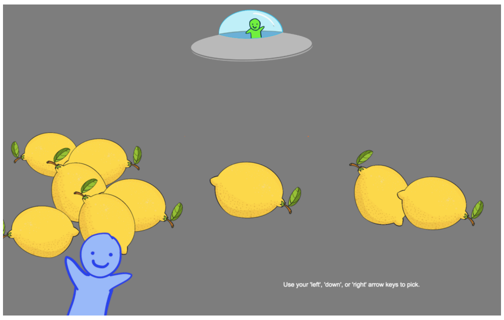


Figure 3: Nominal number validation phase.

Before the experiment continued to verbal number agreement, participants first underwent 2 basic attention checks. (In order to measure attention throughout the entire experiment, 2 more such items occurred after the verbal test phase near the end of the experiment.) As in the English word discrimination item in the sound check preceding the main experiment, participants were presented with an audio clip and asked to discriminate the word they had heard from four orthographic choices arranged on the screen by making a selection with their arrow keys. Rather than monosyllabic English words, the audio clip and the on-screen choices were well-formed (i.e. inflected) forms from the artificial language. The four options were created by crossing two stems with two appropriate (i.e. verbal vs. nominal) suffixes. Given the additional complexity of these items (artificial language rather than English words, bisyllabic rather than monosyllabic, sharing one syllable each with two incorrect choices displayed), the participants only had to make the correct choice in *2 or more* of these 4 attention check items for their data to be included in analysis.

2.3.2 Verbal number

The verbal training and test phases mirrored the nominal ones closely. Participants were told that Nika, the spaceship-piloting character, was going to make a smoothie, either collecting or discarding ingredients with their flying saucer. Laru would instruct Nika based on a recipe on their computer.

Figure 4 represents the verbal training phase, which also began with an exposure portion. Participants were told that Laru was first going to tell Nika to beam things up, then heard an audio clip of an imperative sentence with verbal agreement with the number of the object, and saw an animation of the corresponding number of fruit moving up the screen into the flying saucer (11 repetitions). Next, participants were told that the computer with the recipe was running out of battery, and so was making weird noises that might drown out part of Laru's helpful instructions. The participants were asked to nonetheless choose between two quantities of fruit based on the audio they heard using their arrow keys, and given feedback on their choice (12 trials). As in the nominal pick-the-right-one part of training, the correct number value and the referent placement on the screen were random. The audio clip consisted of an imperative verb inflected to agree with a nominal object of the correct number, whereas the fully disambiguating number-inflected noun was blocked out by a synthesized pulse train noise. As a result, though one of the verbal number agreement suffixes mapped to only one number value (i.e. to singular in ABB, to plural in AAB, and to dual in ABA), the other was syncretic between two values. Therefore the displayed pairs of number values {sg/du, sg/pl, du/pl} excluded the pair that was either ambiguous given the syncretic form or incompatible with the dedicated one (i.e. du/pl in ABB, sg/du in AAB, and sg/pl in ABA). They were told that Nika's spaceship would show them the right answer, even if they made the wrong choice. Feedback consisted of the correct referent being animated to get beamed up or thrown away. Exposure items and pick-the-right-one training items alternated in 3 blocks of 3 and 4, respectively.

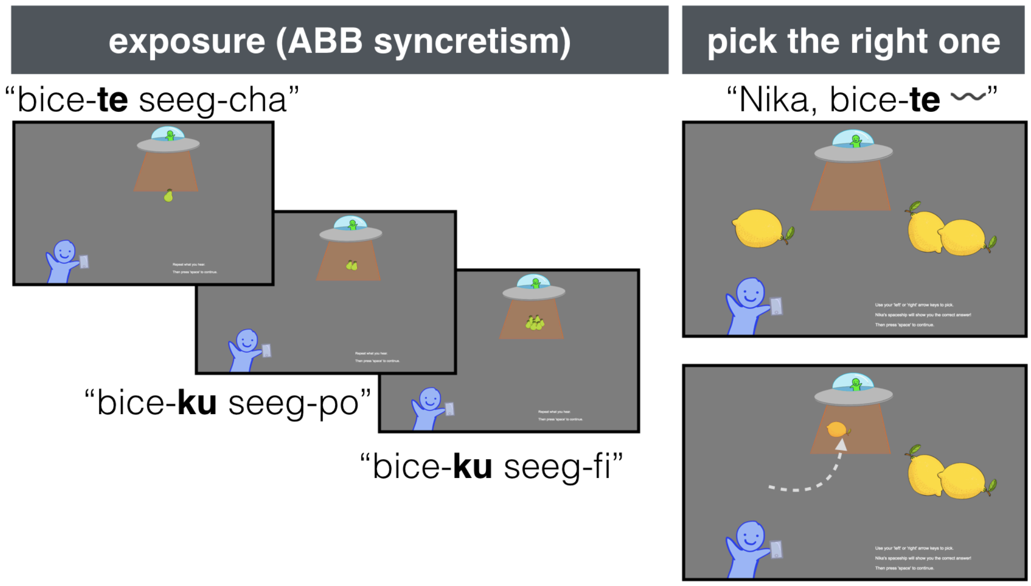


Figure 4: Verbal number agreement training phase.

The crucial test items of this experiment measured learning of the different syncretisms via a forced-choice referent selection task between two quantities of fruit, based only on imperative verb forms bearing object number agreement, with the fully disambiguating noun obscured by noise, with no feedback. As with the pick-the-right-one portion of the verbal training phase, one pair of displayed number values is incompatible with the dedicated verbal form (i.e. du/pl with A in ABB, sg/du with B in AAB, and sg/pl with A in ABA) and is therefore excluded with the audio. However, unlike the training items, the training phase included presentations of the ambiguous pair with the syncretic form. This gives 5 possible items, 4 of which are crucial to measuring learning of the syncretism, which were repeated 4 times (with different fruit stems — apples, bananas, oranges, watermelons) for a total of 20 total test items (of which 16 are crucial). Each participant saw these 20 referent selection items in a randomized order, and the placement of the correct referent (if there was only one) was counterbalanced between the left and the right of the screen. No feedback was given on their choices, and so these items can be represented by the top right screenshot in Figure 4.

Finally,[9] participants were thanked and shown an animation of the completed smoothie blending up. Upon completion of the experiment on Pavlovia, participants returned to the Amazon Mechanical Turk page and answered two additional questions, given in Figure 5.

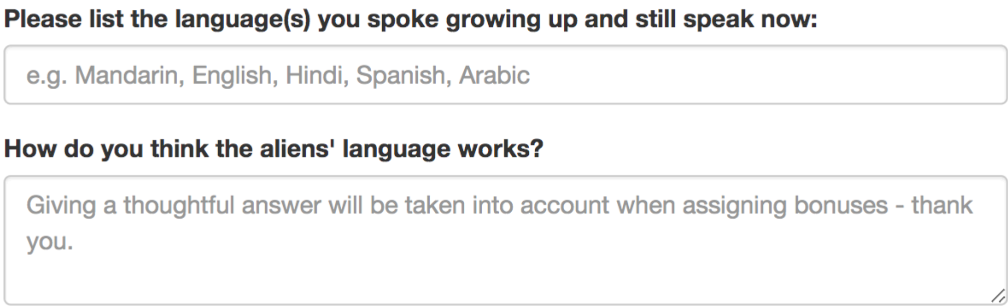


Figure 5: Exit questions.

2.4 *Participants*

Participants were recruited from Amazon Mechanical Turk workers, whose locations were set to the United States, with a good track record on other Amazon Mechanical Turk tasks ($>$ 95%, $>$ 50 total approved). Though 325 English-speaking workers were recruited (ABB: 108, AAB: 108, ABA: 109), only 257 of these actually completed the task on Pavlovia (ABB: 82, AAB: 85, ABA: 90) and were paid USD $3.50; the rest had missing or incomplete data files and were paid USD $1.75.

Of those 257 participants who completed the entire experiment, 149 participants (ABB: 52, AAB: 49, ABA: 48) responded accurately on both basic sound checks *and* responded accurately on two or more attention checks. All but 3 self-reported in the post-experiment language background question that they spoke English growing up. Some of these 149 participants were also given a bonus, bringing their compensation up to USD $4.00 total, for insightful descriptions of the alien language that demonstrated clear engagement with the patterns.

3 Results

Accuracy on the referent selection items, for both the three-way number distinction and the crucial verbal number agreement syncretisms, is a binary measure that is analyzed using logistic mixed-effects models with grammar condition as the main predictor. As both the cross-classifying and linear hierarchy theories of number discussed in Section 1 predict that the dual-plural syncretism in the ABB condition should be learnable, the grammar condition is treatment coded with that ABB grammar as the baseline. As a result, the other grammar conditions, whose relative learnabilities are precisely at issue in comparing these theories, are compared to ABB as a fixed level. All models also include a random intercept for each participant, and were fit in R (R Core Team 2020) using the *glmer* function in the *lmer4* package, which provides p-values based on asymptotic Wald tests (Bates et al. 2015: 101). These statistical analyses follow Maldonado & Culbertson (2020).

[9] The experiment included an exploratory block of 8 artificial language judgment items after all other blocks, to investigate whether a single task could capture both grammaticality and truth value judgments. A new (human) character provided attempts at describing a scene in the alien language, with utterances that varied on whether the nominal number affix accurately or inaccurately reflected the referent amount, and whether the verbal number agreement with that nominal affix was grammatical or ungrammatical. Participants were asked whether the character had “said it wrong” or “said it right”. This binary choice proved not to provide enough insight into participants’ grasp of the grammar. Since these items occurred after all crucial items, they leave no influence on the other results. Further discussion of these items is therefore not included.

3.1 *Learning a three-way number distinction: nominal number*

Figure 6 shows participants' mean accuracy on the three-way forced choice referent selection items in the nominal number ending validation phase. Means for each grammar condition are represented by larger black dots, with error bars showing standard error on by-participant means. Smaller gray dots show each individual participant's mean accuracy across the 8 items they saw. The dotted line marks chance value (0.33 for forced choice with three options).

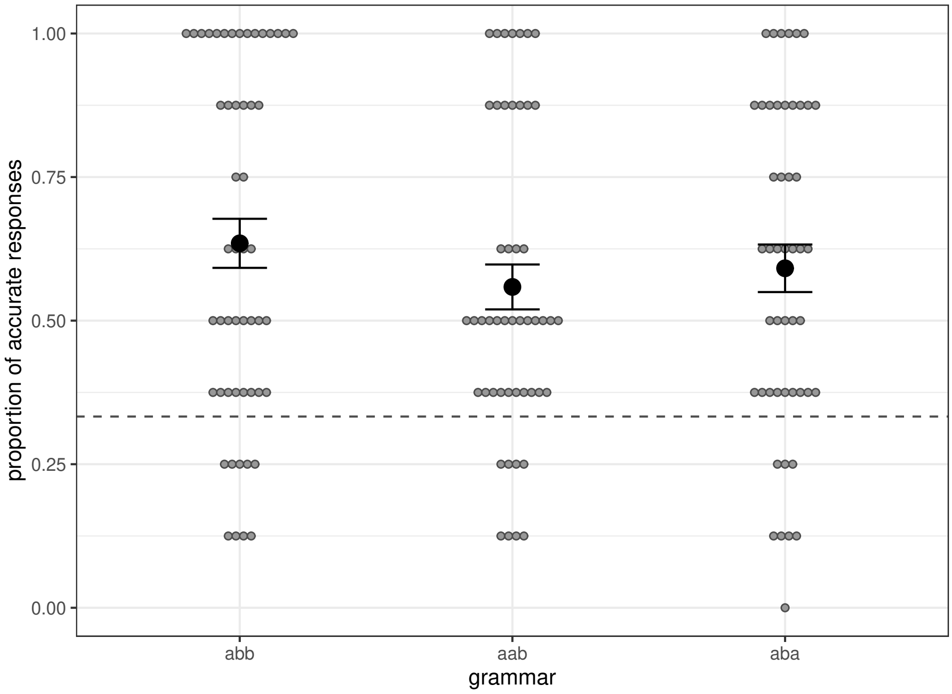


Figure 6: Participant accuracy on nominal endings by grammar.

As all three grammars are identical in the nominal domain, and participants in all conditions have had identical experiences at this point, no difference in nominal referent selection accuracy by grammar is expected. Indeed, Figure 6 shows that mean participant accuracy is notably above chance in all grammars. Additionally, a mixed logit model with random intercepts for each participant, whose output is given in Table 9, confirms that participants across all three grammars learned the three-way distinction. Accuracy in the ABB condition was significantly above chance (intercept: $p < 0.001$***), and neither AAB nor ABA was significantly different from that ABB baseline ($p = 0.147$ and 0.336, respectively).

Table 9: Output of logit model for nominal number validation items.

Formula: validation.corr ~ grammar + (1 I participant)					
Fixed effects	Estimate	Std. Error	z-value	p-value	
(Intercept)	0.8173	0.2269	3.602	0.000316	***
grammar:aab	−0.4623	0.3190	−1.449	0.147293	n.s.
grammar:aba	−0.3093	0.3217	−0.961	0.336316	n.s.

Participants were accurate across all number values: their accuracy on dual-marked items of 0.619 was significantly above chance (intercept: $p < 0.001$***), with no significant difference from that dual baseline for either singular (mean = 0.573, $p = 0.183$) or plural items (mean = 0.595, $p = 0.447$).

3.2 *Learning different number syncretisms: verbal number agreement*

Figure 7 shows participants' mean accuracy on the 16 *crucial* items in the verbal number agreement test phase. Though the test phase included an item type with no wrong answer, where the syncretic verbal marker appears with precisely the number values that it neutralizes, this kind of item is excluded from the visualizations and analysis.

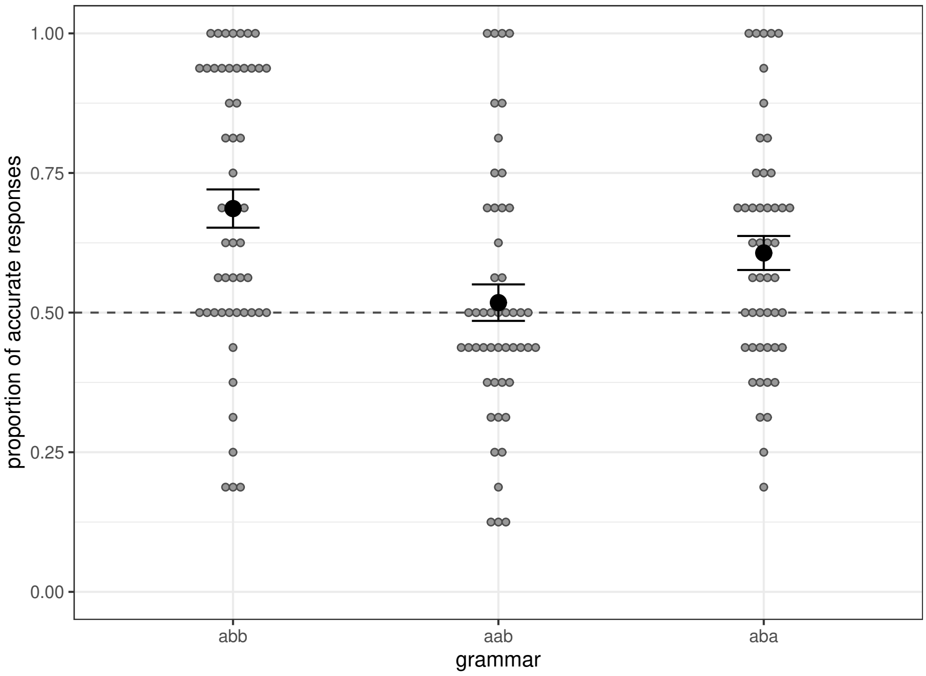


Figure 7: Participant accuracy on verbal endings by grammar.

Whereas the three-way referent selection for nominal forms represented in Figure 6 had a 0.33 chance level, the two-way forced choice items for the verbal test phase had a chance value of 0.5, represented by the dotted line in Figure 7. Participants in the dual-plural (ABB) grammar condition performed visibly well above chance, as predicted by all theories considered. The significant model intercept of a treatment coded mixed logit model (intercept: $p < 0.001$***) confirms that the ABB condition's participants' visibly high average accuracy of 0.686 (with standard error of 0.034) is indeed significantly above chance. The average accuracy of participants in both the singular-dual (AAB) and the singular-plural (ABA) grammar conditions was lower than those in ABB. In particular, many participants in the AAB condition seem to be performing around chance (mean = 0.518, se = 0.033). The model confirms that their accuracy was strongly significantly lower than that baseline ($p < 0.001$***) (output in Table 10). Conversely, although participants in the ABA condition had a lower average accuracy (mean = 0.607, se = 0.030) than the ABB participants, this difference is marginally not significant in the logit model ($p = 0.0567$).

Table 10: Output of logit model for crucial verbal test items.

Formula: trial.corr ∼ grammar + (1 | participant)					
Fixed effects	Estimate	Std. Error	z-value	p-value	
(Intercept)	1.0304	0.1770	5.823	5.78e−09	***
grammar:aab	−0.9077	0.2484	−3.654	0.000258	***
grammar:aba	−0.4765	0.2500	−1.906	0.056668	n.s.

4 Discussion

This experiment aimed to test whether learners make use of features to represent and group number values. To do so, it presented learners with morphological patterns of syncretism that target different groupings of those number values. The results confirm that adult learners are (a) sensitive to the relevant (non-native) number values, (b) able to learn syncretism patterns neutralizing those values in an artificial language setting, and (c) specifically better at *feature-based* syncretism patterns. Learners were not equally successful at all possible groupings of number values. Rather, they were more successful when the morphological pattern they learned involves number values that form a feature-based natural class. Specifically, the more easily learned syncretisms are those that *linearly hierarchical features* are able to represent. I discuss each of these findings in turn, before treating typological data and possible alternative interpretations of the results.

4.1 *Number morphology is learnable in artificial language settings*

First of all, the patterns were based on non-native number contrasts between singular, dual, and plural, so it is crucial for participants to be sensitive to them. This is especially true as adults appear to reduce their sensitivity to certain kinds of non-native contrasts (particularly in phonology: see Werker & Tees 1984; Werker 1994, i.a.). It is therefore important for artificial language learning experiments like this one to demonstrate participants' sensitivity to the specific domain under investigation. This experiment verified learners' abilities in the domain of number marking by exposing them to a singular-dual-plural number contrast on nouns, then testing their comprehension of those forms in a referent selection task.

Results show that participants succeeded in accurately picking referent quantities based on the three singular, dual, and plural nominal endings equally well, in all grammars. Their high mean accuracy, which was significantly above chance levels ($p < 0.001$***), and the lack of significant differences between number values, or grammar conditions, confirm that the dual is a non-native contrast that is learnable by adults in an artificial language setting. Written responses completed after the experiment corroborate participants' sensitivity to the dual suffix as a number marker that contrasts with singular and plural. Representative quotes from two English-speaking monolingual participants are given here: "The suffix of each word determines the number of items you're trying to describe. *Po refers to two items*, cha to one item and fi for several item[s]," and, "It looks like the endings of the words denote quantities. Cha = 1, Fi = many, *po = 2*" [emphasis in both quotes mine]. This finding in number marking specifically adds to the growing evidence that adults are sensitive to non-native phi-featural contrasts in number and person, as demonstrated in Finley & Wiemers (2013) and Maldonado & Culbertson (2020).

4.2 *Features are active in learning of number morphology patterns*

The three grammar conditions in this experiment grouped different pairs from the three number values — singular, dual, and plural — into an agreement syncretism: dual-plural (ABB), singular-dual (AAB), and singular-plural (ABA). If there is no further abstraction to the three number values, then any of those three sets should be equally easy or difficult to learn. If, on the other hand, more abstract features underlie the three number values and group them into natural classes, then particular groupings of singular, dual, and plural should be more readily learned than others. Specifically, all feature-based approaches to singular, dual, and plural number represent dual and plural as a morphosyntactically relevant natural class, though the theories make opposing predictions for the natural class membership of the singular. As a result, a feature-based approach to number predicts that learners in the ABB condition will perform significantly better than chance. Further, one of the two other grammars will not be significantly different, whereas the last grammar will be significantly worse. If learners are using cross-classifying features, such as those of Harbour, then the singular-plural (ABA) grammar should be the significantly worse condition. If they are using linearly dependent features, such as those of Smith et al. (2019), then it should be the singular-dual (AAB) grammar condition that is significantly worse.

Indeed, while participants performed very similarly across the board on learning the nominal dual, they differed significantly on learning patterns that grouped the three numbers in different ways. As predicted by the feature-based approach (and influence from participants' experience with English — see Section 4.4.1), learners in the ABB condition were the most successful at picking referent quantities based only on the syncretic verbal agreement forms. Learners in the ABA condition were not significantly different than those in the ABB baseline, but learners in the AAB condition were significantly worse. This differential performance suggests that learning of morphological patterns of number marking is mediated by feature representations. This follows a rich literature in phonology that shows that learners more easily learn phonological patterns that can be described using a single feature than those involving a featurally random set of sounds (Saffran & Thiessen 2003; Cristià & Seidl 2008; Finley & Badecker 2009; Cristià et al. 2013; Gallagher 2013; Linzen & Gallagher 2014, i.a. surveyed in Culbertson & Schuler 2019). It directly contributes to nascent analogous findings that learners prefer feature-based generalizations about morphological values, as well (Finley & Wiemers 2015: Exp 2; Maldonado & Culbertson 2020).

4.2.1 By-condition results are consistent with use of linearly hierarchical features

In particular, the differential performance between the more readily learnable ABB and ABA syncretisms on the one hand, and the less learnable AAB syncretism on the other, suggests that participants are specifically using *linearly organized* features in learning number morphology. Recall that we cast Smith et al. (2019) in Nanosyntax, a theory that takes linear feature hierarchy to be a core architectural tenet; the resulting illustrative system equips learners with a universal linear hierarchy of privative features for building number representations. Learning number morphology is accomplished by finding lexical representations that perfectly match or encompass (as a superset) the featural heads in a given number value. The predictions of this expository linear feature hierarchy theory are most compatible with how participants performed throughout the experiment.

First, these linear features predict success in learning ABB syncretisms that neutralize dual and plural. Because the dual contains all of the featural material that the plural does, both can be pronounced using the same lexical representation that contains the union of their features. The singular is represented with even less material than either the dual or the plural, so it can be spelled out with a separate lexical representation. A learner using Smith et al. (2019)-style features can posit those two lexical entries (given in (11)) and accurately learn the grammar in the ABB syncretism condition. This is consistent with the high accuracy (0.686, significantly better than chance ($p < 0.001$***)) that participants learning the ABB dual-plural syncretism had when using the syncretic verbal endings to select the correct quantity of referent.

On the other hand, this linear feature hierarchy cannot represent AAB syncretisms that neutralize singular and dual, and predict difficulty in learning such a pattern. Because Smith et al.'s (2019) structure for plural intervenes between their structures for singular and dual, any maximal representation that could match both singular and dual (such as the bottom lexical entry given in (13)) would *necessarily* also match plural, and accidentally derive a singular-plural ABA syncretism instead. One would need a completely differently ordered linear feature hierarchy to derive an AAB pattern; as mentioned in Section 1.4.2, Smith et al. admit cross-linguistic variation in the topmost feature in the linear hierarchy, and suggest that languages containing AAB lexical noun suppletion have parametrically chosen a different (pl > du > sg) hierarchy (2019: 1072–4). But learners with a single universal hierarchical ordering of features (du > pl > sg) and those Nanosyntactic spellout assumptions, could only "mimic a surface [AAB] pattern" by learning three separate lexical representations, one for each number agreement value, with accidental homophony for the singular and dual (ibid. fn7). "Standard working assumptions [for] morphologists" treat truly accidental homophony as to be avoided in analysis and dispreferred in learning (Halle & Marantz 2008: 56; Müller 2004; 2005; Baerman et al. 2005; Pertsova 2007; 2011; Albright & Fuß 2012; Bobaljik 2012: 35; Bank 2017). Such a dispreference against homophony may be violable, but learners will take longer or require more types of evidence to overcome it. In the context of this experiment, any of these interpretations is consistent with how participants learning the AAB singular-dual syncretism appear to be no better than chance (accuracy = 0.518) at referent selection based on syncretic verbal agreement, and do so significantly worse than the ABB-learning participants ($p < 0.001$***).

Conversely, as already described, a linear feature hierarchy *can* represent ABA syncretisms. A learner equipped with Smith et al. (2019)-style features can posit the two lexical entries given in (13) and neutralize singular and plural verbal agreement. Though learners of this ABA singular-plural syncretism are slightly less accurate on average (0.607) than the ABB learners, this difference is marginally not significant ($p = 0.0567$); this accords with how both ABB and ABA syncretisms are equally generable in this theory.

Learners' performance on the three different verbal number agreement syncretism conditions thus matches the predictions made by Smith et al. (2019)-style linear number features.

4.2.2 By-condition results are less consistent with cross-classifying features

Although the success of the ABB condition participants is consistent with abstract feature theories of number in general, results in the other grammar conditions are harder to square with cross-classifying feature representations in particular. Recall that our representative cross-classifying theory, Harbour (2014; 2016), posits an innate set of binary morphosemantic features that may be activated in specific languages. The meanings of singular, dual, and plural number are generated by bundling different values of two such innately available morphosemantic binary features: [±atomic] and [±minimal]. Vocabulary Items may match a subset of the valued features in those bundles. When bundles have a valued feature in common, as dual and plural do with [−atomic], or as singular and dual do with [+minimal], then learners should be able to posit featurally underspecified Vocabulary Items to neutralize those bundles. It should be equally easy to hypothesize the Vocabulary Items in (3)d and (3)e to capture the ABB syncretism grammar, and to hypothesize the Vocabulary Items in (5)d and (5)e to learn the AAB syncretism grammar. The results do not support this prediction: ABB syncretism learners were indeed significantly better than chance at referent selection ($p < 0.001$***), but AAB syncretism learners were significantly worse ($p < 0.001$***) than their ABB counterparts. The feature architecture of Harbour's number theory offers no explanation for this asymmetry.

Harbour's feature bundles for singular and plural number (agreement) have no value of [±atomic] or [±minimal] in common. No Vocabulary Item bearing one of those four possible values can be inserted in both those contexts. This would seem to rule out singular-plural syncretisms within Harbour's number theory, contra the actual experimental results: participants found the ABA syncretism no less learnable than the ABB syncretism ($p = 0.0567$). But what if a Vocabulary Item were even more radically underspecified, and bore no valued number feature at all? A default Vocabulary Item specified only for its categorial status — in this case, of being verbal number agreement — *could* target singular agreement and plural agreement despite their being "feature bundles that have no [valued] features in common at all" (Kramer 2016: 99). A learner could posit a fully specified Vocabulary Item for the dual alongside this extremely underspecified Vocabulary Item and derive an apparent ABA pattern via default exponence. Admitting default exponence (or accidental homophony) in this way could make the success of ABA condition participants comprehensible within the Harbour-style number feature theory. However, this would make the difficulty of the AAB condition even harder to explain in this framework; if an apparent singular-plural ABA neutralization could be captured by a specific dual exponent and a default exponent, so too should the AAB pattern (with a specific plural exponent and a default exponent). This means the theory at its most permissive should allow at least two ways of deriving the ABB and AAB patterns, along with at least one way of deriving the ABA pattern with default exponence. This makes the participants' singling out of the AAB pattern particularly difficult to explain with Harbour-style cross-classifying number features.

4.3 Typology does not fully mirror experimentally preferred neutralizations

This behavioral evidence for feature-based natural classes represents an important addition to linguists' toolbox for theorizing about innate features and natural classes. Traditional investigations have looked to typological surveys of morphological patterns to substantiate the existence of natural classes. However, a multitude of factors may intervene between "the generative capacity of the linguistic system...and particular features of [a set of] language[s]" (Maldonado & Culbertson 2020: 6), from historical contingencies, to pressures from language's communicative function.

Indeed, the attestation patterns for syncretisms and suppletions involving singular, dual, and plural number *do not* appear to directly replicate the learning results presented here. Table 11 summarizes my survey of 30 spoken languages with a singular-dual-plural number system for any morphological neutralization patterns anywhere number is marked in the grammar (Appendix). This includes pronominal stems and number affixes, nominal number affixes, verbal number agreement affixes, and lexical stems that supplete for number.

Table 11: sg-du-pl languages with number neutralization patterns (n = 30).

	ABB du-pl	AAB sg-du	ABA sg-pl
syncretism	22	6	1
suppletion	14	5	1
total	26	10	2?

My typological survey reaffirms the prevalence of ABB dual-plural syncretisms and supletions, finding them in 26 of the 30 considered languages, in multiple domains. For example, the Hopi example in (17) demonstrates an ABB pattern in the third person demonstrative pronoun, while the bottom of Table 12 shows Koasati verbs whose dual and plural forms share a suppletive stem. The typological dominance of this neutralization is paralleled by participants' greatest success with the experimental ABB grammar.

(17) *Hopi*: ABB pronouns, AAB verb suppletion

	a.	b.	c.
	pam	puma	puma
	that.**SG**	that.**DU**/PL	that.DU/**PL**
	wari	wari	yu'tu
	run.**SG**/DU	run.SG/**DU**	run.**PL**
	he/she ran.	they (2) ran.	they (>2) ran.

Table 12: *Koasati*: ABB & AAB suppletive verbs (Kimball 1985: 273–275, Table 10.3, fn12).

	verb	singular	dual	plural
ABC	'to stand'	haccá:lin	hikkí:lin	lokkó:lin
	'to sit'	cokkó:lin	cikkí:kan	í:san
	'to dwell'	á:tan	áswan	í:san
AAB	'to die'	íllin		hápkan
	'to go about'	á:yan		yomáhlin
	'to go'	aɬí:yan		amá:kan
	'to come'	óntin		ilmá:kan
ABB	'to be lost'	nak aɬɬan	wasátkan	
	'to run'	walí:kan	tóɬkan	
	'to clamber up'	onno-halí:kan	onno-káhkan	
	'to exit'	ac-halí:kan	as-káhkan	
	'to enter'	cok-halí:kan	cok-káhkan	
	'to release' (obj)	acapɬ́kan	askáhlin	
	'to pick up' (obj)	í:sin	píhlin	
	'to hit' (pluractional)	batáplin	bóklin	
	'to make noise' (pluractional)	naksá:kan	sakáplin	

I also find AAB singular-dual neutralization patterns in 10 languages, such as in Zuni's intransitive verbal agreement prefix (18). Intriguingly, all 10 languages with an AAB pattern also demonstrate an ABB dual-plural pattern elsewhere; alongside their ABB patterns, both Hopi (17) and Koasati (Table 12) have verbs whose singular and dual forms share a suppletive stem.

(18) *Zuni*: AAB verbal agreement prefix (Bunzel 1933: 423, 492)

	a.	b.	c.
	Ø-kʷato-k̯ä	a:tci Ø-kʷato-k̯ä	ho'i tem'ɬ u-kʷato-k̯ä
	SG/DU-enter-PST	3.DU SG/DU-enter-PST	people all PL-enter-PST
	S/he went in.	They (both) went in.	All the people went in.

This relative prevalence of singular-dual AAB patterns in the typology is surprising, given participants' significant learning dispreference for the experimental AAB grammar. One might be tempted to discount some of the AAB patterns, especially syncretisms where the shared singular and dual form is null, as in (18). However, if a Nanosyntactic interpretation of Section 4.2.1's linear-hierarchy-based analysis of the results is on the right track, analyzing such patterns as a marked plural with a default is not an option: Nanosyntax's Superset Principle rules out default exponents. Rather, as previously discussed, the only theory-internal way to create the appearance of a non-contiguous neutralization like the singular-dual AAB pattern is two separate, homophonous lexical entries. Within the experiment, limited training data may not be sufficient to override antihomophony biases — but perhaps in natural language, morphological paradigms are small enough and their forms abundantly represented enough in children's input to violate antihomophony despite underlying feature-driven constraints (Harbour 2016: 15). Such a "surfeit of the stimulus" might lead to wider representation of AAB in the typology than otherwise expected (ibid.).

Finally, my typological survey reveals a lack of robust ABA patterns. I find only 2 tentative cases in 30 languages. For example, Yagua's 2nd person free/suffixal pronoun stems (given in Table 13) are "by no means a robust" example of ABA (Smith et al. 2019: 1065). First, the free/suffixal 2nd person dual pronoun stem resembles the prefixal 3rd person singular pronoun stem. Moreover, Smith et al. (2019) identify an ABB pattern in the 1st exclusive, an ABA pattern in the 2nd person, and an AAB pattern in the 3rd person (online supplemental materials, Yagua). No number feature theory derives three overlapping natural classes, so if Yagua's pronouns really do exhibit ABB, ABA, *and* AAB patterns all in a single language, their exponence cannot be captured via feature-based theories.

Table 13: Yagua: ABA(?) pronoun stems (Payne 1985: 41–44, Tables 2.1–2.2).

	free/suffixal pronouns				prefixal pronouns			
	singular	dual	plural		singular	dual	plural	
1st excl	ráy	nááy	núúy	ABB?	ráy	nááy	núúy	ABB?
1st incl	–	vúúy			–	vúúy		
2nd	jiy	saadá	jiryéy	ABA?	jíy	saaná	jíryey	ABA?
3rd	níí	naadá	ríy	AAB?	sa	naada	riy	

This paucity of singular-plural ABA patterns in the typology is surprising given participants' accurate performance on the experimental ABA grammar. Perhaps singular quantities and plural quantities are functionally important to keep distinct in real situations, to the detriment of such neutralizations.[10] The inverse of that neutralized form, a marked dual, may not be particularly useful either: it may not be pragmatically necessary to linguistically mark natural pairs, for example. ABA patterns may have low functionality (particularly compared to AAB patterns, which might have the same communicative function as a (limited) paucal number). Over time, these pressures might lead to less accurate diachronic replication and sparser typological representation of ABA patterns than otherwise expected.

This partial match between learnability and typology resembles experimental findings in much of the phonological and morphosyntactic artificial language learning literature. Though learner preferences often parallel the strongest typological asymmetries (see Culbertson 2012; Moreton & Pater 2012a for reviews), it is quite common to find mismatches of various kinds (Moreton & Pater 2012b), often within the same experiment. First of all, adult and child learners are often indifferent between patterns that have very different rates of attestation (as the ABB and ABA neutralizations do): there may be no significant difference at all (Finley & Badecker 2009: Exp 3; see Sections 3 & 4 of Moreton & Pater 2012b for several other phonological examples; Tabullo et al. 2012; Maldonado & Culbertson 2020: Exp 3 (condition 1); Culbertson et al. 2020: Experiments 1 & 2 (Hebrew,

[10] Thanks to Kenny Smith for suggesting this idea to me.

condition 2 vs. 3 & 4)) or only a marginal one (Culbertson et al. 2012; Culbertson et al. 2020: Exp 1 (Hebrew, conditions 1 & 2 vs. 3) and Exp 2 (French, condition 1 vs. 4)). For example, Tabullo et al. (2012) find that participants exposed to an artificial language displaying the extremely rare OSV word order (0.5% of world languages) are just as accurate at an error detection comprehension task as those who learned the much more prevalent SOV (45%) or SVO (42%) word orders. Conversely, learners may show preferences between patterns that are equally (ill-)attested (Moreton 2008: Exp 2) or even completely unattested in real language (Maldonado & Culbertson 2020: Exp 2 (Second-inclusive vs. Third-inclusive)). Finally, artificial language learners may prefer patterns in reverse of relative typological prevalence (Koo 2007: Ch. 2 (vowel vs. consonant identity); Tabullo et al. 2012; Maldonado & Culbertson 2020: Exp 1B, condition 2 vs. 3; Maldonado et al. 2020: prefix condition; Culbertson et al. 2020: Exp 1 (French, condition 1 vs. 3)) or implicational universals (Cook 1988: condition D: VO ⇒ Prep-N). For example, person prefixes precede number prefixes more often (8 languages) than the inverse order (only 1 language), but Maldonado et al. (2020) find that learners preferred the latter pattern anyway, just as AAB neutralizations typologically outweigh ABA patterns, but are less learnable experimentally.

This experiment's mismatch between learners' preferences among the three possible neutralizations and the relative prevalence of those neutralizations in the world's languages is therefore instructive. It suggests that feature-based learning biases do not necessarily translate directly into typology, and — like these previous studies — encourages further investigation into what factors may disrupt that relationship. In this domain of singular, dual, and plural number neutralizations, it is possible that communicative functional pressures only weakly at play in this experiment could override feature-based preferences against AAB patterns and for ABA patterns.

4.4 *Alternative interpretations of the results to consider*

The foregoing discussion has argued that the results of this artificial language learning experiment are best understood as support for innate number features. But English conflates the meaning of the dual into the plural throughout its grammar: "two dog-s" has the same morphological marking as "three dog-s." The ABB syncretism neutralizes these same values, and was (marginally) the easiest grammar to learn in this experiment. Could influence from participants' English number system be partially responsible for these learning results, too?

The relationship between adult native language grammars and artificial language learning tasks has been explored in the main previous artificial language study on morphological feature organization: Maldonado & Culbertson (2020) simply expected that "native-like paradigms [would] be easiest to learn" (18). But what makes an artificial language pattern native-like? I first consider surface similarity to and difference from English number morphology as a factor in each grammar condition. I then evaluate possible feature-based representational limitations that native English lexical entries in each theory might transfer to learning new systems (Hawkins & Chan 1997; Tsimpli & Dimitrakopoulou 2007; i.a.).

4.4.1 Surface comparison to English number morphology

Perhaps participants are not analyzing the featural content of the artificial language forms at all. Could their differential performance across conditions stem solely from stubbornly comparing all artificial language conditions to English-like Vocabulary Items, which conflate the dual and the plural?

Such a strategy would most obviously privilege the ABB verbal syncretism in learning. An English-speaking participant could map the English dual-plural conflation directly to the artificial language's dual-plural syncretism in the verbal domain. This surface similarity between English and the ABB grammar condition might have contributed to participants' learning it marginally better (ABB accuracy = 0.686, se = 0.034) than the next easiest grammar (ABA accuracy = 0.607, se = 0.030), albeit non-significantly so ($p = 0.0567$). English influence might explain this observed marginal difference between ABB and ABA, which should be equally learnable from a solely theoretical standpoint.

Although ABB patterns are the most straightforwardly English-like, direct surface comparison to English could conceivably have privileged the ABA grammar condition as well. An English-speaking participant might not have any English analogue to map to the artificial language's singular-plural syncretism, but the surface dissimilarity between English and the non-syncretic dual markers might make them even more salient and learnable to them. This focused salient difference between English and the ABA grammar condition might have contributed to participants' learning it better (ABA accuracy = 0.607, se = 0.030) than the completely non-English-like AAB grammar (accuracy = 0.518, se = 0.033).

However, participants' behavior on nominal items seems to counter the overall idea that the dual is only evaluated in relation to the English number system. If participants were indeed rigidly imposing English's singular-plural system on the artificial languages, we might expect them to behave exceptionally when faced with the separate marking of nouns for groups of two. They might incorrectly collapse dual referents into the meaning of the plural (a sort of ABB effect). Alternatively, they might be particularly attentive and accurate on dual items (the proposed ABA effect). But participants are not significantly more or less accurate on picking the correct referent quantity based on a dual-marked nominal than they are with either other number. This is true across all conditions (vs. singular $p = 0.183$, vs. plural $p = 0.447$).

Direct surface comparison with English thus does not seem to fully explain participants' behavior across nominal and verbal items, but may still have played a role in how easily they learned the ABB, and possibly the ABA, syncretism patterns.

4.4.2 Direct translations from English number morphology

Could the influence from participants' native language be less holistic, and more based on direct translation of English forms? Perhaps participants calqued the artificial language forms to existing English lexical entries, with a concomitant boost to accuracy at referent selection when hearing those items.

In fact, neither theory proposes structures for the dual-plural conflation of English that match the crucial verbal agreement syncretisms. With a Nanosyntax-style linear feature hierarchy, the superset lexical entry that neutralizes dual and plural verbal number agreement (given in (11)) includes a verbalizing head in order to distinguish dual and plural exponence in a verbal agreement complex from number marking on nouns, so it has no correlate in the English lexicon. Similarly, with DM-style cross-classifying feature bundles, the artificial language VIs (given in (3)d and (3)e) include an additional contextual feature, making their insertion contexts distinct from the English ones. Direct translation between identical featural structures (trees or bundles) in the English lexicon and the artificial languages therefore cannot explain differences between grammar conditions.

One theory *does* actually posit an identical structure in English and in the artificial languages: the Smith et al. (2019)-style representation of the English nominal "plural" and the artificial languages' nominal dual. Nanosyntax takes the linear functional sequence of number features to be universal, meaning that even English *syntactically* represents dual number.[11] English lacks a *morphological* distinction for the dual simply because a single large lexical subtree, given in (9), matches both the dual "two dog-s" and the plural in "three dog-s." The very same tree conflates dual and plural by the Superset Principle in English, and spells out dual nominal number exactly in all the artificial languages. If identity in structure between existing English lexical items and artificial language lexical items were helpful, participants might be expected to perform better on the nominal referent selection items with dual marking than those with plural marking. But, again, mean accuracy on dual items (0.619) and plural items (0.595) is not significantly different ($p = 0.447$). Evidently, participants do not get a boost simply from having exactly identical analogues to English lexical items in the artificial languages. This, and the discussion in the previous subsection, suggest that the precise link between the morphological contrasts of participants' native language and any advantage conferred to similarly structured patterns in artificial grammars is more complex than the straight line drawn by Maldonado & Culbertson (2020).

4.4.3 Interpreting number morphology as numerals

One final possibility is that participants are not tapping into number feature representations as they learn and use the artificial language patterns, but rather taking the number exponents to be postnominal numerals. A numeral interpretation may be reconciliable with the singular and dual number exponents, as discussed above in the context of the direct English imposition strategy. However, the plural number exponent appears with referents in multiple different amounts, ranging from six to nine. Even if participants really were counting the amounts of fruit that appeared with plural number or number agreement markers, there would be no single numeral that would be consistent with all the training input. Moreover, none of the participants who gave a written response to the exit questionnaire mentioned any specific cardinalities other than one and two. Those who referenced the meaning of the plural number exponents described it as "plural", "many", "more than two", "at least three", "multiples", "a bunch", and/or "a group of some". The range of quantities used to train participants on plural marking, as well as the corresponding complete absence of numeral interpretations for the plural, suggests participants really were treating the affixes as marking number, not numerals.

5 Conclusion

Number systems differentiating singular, dual, and plural have been treated by multiple different theories of morphosyntactic organization. These theories make differing predictions about which groups of values may be represented as natural classes and thereby systematically neutralized by the grammar. This paper has tested these predictions experimentally, by training and testing participants on artificial languages differing in which

[11] This is also proposed in Section 4.5.3 of Aravind's (2018) dissertation, which explores the idea that all languages, "even those that lack the relevant morpho-syntactic correlates, [are] furnished with [the] feature[s for] dual" (138). The English words 'both' and 'all' reveal the dual-plural distinction (when combined with a universal quantifier), but everywhere else (in number marking on nominals and verbal agreement), the dual is syncretic with the plural.

pairs of number values participate in syncretism. Results show that participants prefer grouping dual with plural, or singular with plural, to grouping singular with dual. This preference for Smith et al. (2019)-style linear hierarchy-based natural classes provides a new source of evidence that number contrasts *are* represented with features.

This experiment provides insight into the choice of featural representation in the domain of number. These results support the value of artificial language learning experiments in investigating questions of representation in the domain of morphosyntax. Though additional work needs to be done to clarify how this behavioral evidence complements and complicates traditional typological investigations of feature structure, this paper represents an advance in applying experimental methods to the theory of morphosyntactic features.

Abbreviations

SG = singular, DU = dual, PL = plural, ABB = dual-plural, AAB = singular-dual, ABA = singular-plural, O = object agreement, VI = Vocabulary Item, DM = Distributed Morphology

Data Availability/Supplementary files

Supplementary file 1: Data file: Referent selection response data for all participants (in csv format).

Supplementary file 2: Appendix. Number syncretism and suppletion patterns in 30 singular-dual-plural languages.

References

Albright, Adam & Eric Fuß. 2012. Syncretism. In Jochen Trommer (ed.), *Morphology and Phonology of Exponence*, 236–288. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199573721.003.0008.

Aravind, Athulya. 2018. *Presuppositions in context*: MIT dissertation.

Baerman, Matthew, Dunstan Brown & Greville G. Corbett. 2005. *The syntax-morphology interface: A study of syncretism*, vol. 109. Cambridge University Press. https://doi.org/10.1017/cbo9780511486234.

Bale, Alan, Michaël Gagnon & Hrayr Khanjian. 2011. On the relationship between morphological and semantic markedness. *Morphology* 21(2). 197–221. https://doi.org/10.1007/s11525-010-9158-1.

Bank, Sebastian. 2017. Segmentation: a remark on the syncretism principle. *Morphology* 27(1). 1–20. https://doi.org/10.1007/s11525-016-9295-2.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01.

Baunaz, Lena, Liliane Haegeman, Karen De Clercq & Eric Lander. 2018. *Exploring nanosyntax*. Oxford University Press. https://doi.org/10.1093/oso/9780190876746.001.0001.

Bobaljik, Jonathan David. 2012. *Universals in comparative morphology: Suppletion, superlatives, and the structure of words*. MIT Press. https://doi.org/10.7551/mitpress/9069.001.0001.

Bobaljik, Jonathan David, Andrew Nevins & Uli Sauerland. 2011. Preface: On the morphosemantics of agreement features. *Morphology* 21(2). 131–140. https://doi.org/10.1007/s11525-010-9155-4.

Boersma, Paul & David Weenink. 2021. Praat: doing phonetics by computer [computer program]. version 6.1.39 http://www.praat.org/.

Bunzel, Ruth L. 1933. Zuni. In Franz Boas (ed.), *Handbook of American Indian languages*, vol. 3, 385–514. Washington: Government Printing Office.

Caha, Pavel. 2019. Nanosyntax: Some key features. https://ling.auf.net/lingbuzz/004437.

Cook, Vivian J. 1988. Language learners' extrapolation of word order in micro-artificial languages. *Language Learning* 38(4). 497–529.

Corbett, Greville G. 2000. *Number*. Cambridge University Press. https://doi.org/10.1017/cbo9781139164344.

Cristià, Alejandrina, Jeff Mielke, Robert Daland & Sharon Peperkamp. 2013. Similarity in the generalization of implicitly learned sound patterns. *Laboratory Phonology* 4(2). 259–285. https://doi.org/10.1515/lp-2013-0010.

Cristià, Alejandrina & Amanda Seidl. 2008. Is infants' learning of sound patterns constrained by phonological features? *Language Learning and Development* 4(3). 203–227. https://doi.org/10.1080/15475440802143109.

Croft, William. 1990. *Typology and universals*. Cambridge University Press.

Culbertson, Jennifer. 2012. Typological universals as reflections of biased learning: Evidence from artificial language learning. *Language and Linguistics Compass* 6(5). 310–329. https://doi.org/https://doi.org/10.1002/lnc3.338.

Culbertson, Jennifer, Julie Franck, Guillaume Braquet, Magda Barrera Navarro & Inbal Arnon. 2020. A learning bias for word order harmony: Evidence from speakers of non-harmonic languages. *Cognition* 204. 1–18. https://doi.org/10.1016/j.cognition.2020.104392.

Culbertson, Jennifer, Annie Gagliardi & Kenny Smith. 2017. Competition between phonological and semantic cues in noun class learning. *Journal of Memory and Language* 92. 343–358. https://doi.org/10.1016/j.jml.2016.08.001.

Culbertson, Jennifer, Hanna Jarvinen, Frances Haggarty & Kenny Smith. 2019. Children's sensitivity to phonological and semantic cues during noun class learning: Evidence for a phonological bias. *Language* 95(2). 268–293. https://doi.org/10.1353/lan.0.0234.

Culbertson, Jennifer & Kathryn Schuler. 2019. Artificial language learning in children. *Annual Review of Linguistics* 5. 353–373. https://doi.org/10.1146/annurev-linguistics-011718-012329.

Culbertson, Jennifer, Paul Smolensky & Géraldine Legendre. 2012. Learning biases predict a word order universal. *Cognition* 122(3). 306–329. https://doi.org/10.1016/j.cognition.2011.10.017.

Davies, Mark. 2009. The 385+ million word corpus of contemporary american english (1990–2008+): Design, architecture, and linguistic insights. *International Journal of Corpus Linguistics* 14(2). 159–190. https://doi.org/10.1075/ijcl.14.2.02dav.

D'Onofrio, Annette. 2014. Phonetic detail and dimensionality in sound-shape correspondences: Refining the bouba-kiki paradigm. *Language and Speech* 57(3). 367–393. https://doi.org/10.1177/0023830913507694.

Embick, David. 2016. On the targets of phonological realization. In Vera Gribanova & Stephanie Shih (eds.), *The Morphosyntax-Phonology Connection: Locality and Directionality at the Interface*, 255–283. Oxford University Press.

Finley, Sara & William Badecker. 2009. Artificial language learning and feature-based generalization. *Journal of Memory and Language* 61(3). 423–437. https://doi.org/10.1016/j.jml.2009.05.002.

Finley, Sara & Elizabeth Wiemers. 2013. Rapid learning of morphological paradigms. In *Proceedings of the Annual Meeting of the Cognitive Science Society*, vol. 35.

Finley, Sara & Elizabeth Wiemers. 2015. Phonological and semantic consistency as cues for learning morphological systems. In *Proceedings of WCCFL 32*. 11–20. Cascadilla Press.

Gallagher, Gillian. 2013. Learning the identity effect as an artificial language: Bias and generalisation. *Phonology* 253–295. https://doi.org/10.1017/s0952675713000134.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H Greenberg (ed.), *Universals of language*, 73–113.

Greenberg, Joseph H. 1966. *Language Universals with Special Reference to Feature Hierarchies*. de Gruyter Mouton.

Hale, Kenneth. 1973. Person marking in Walbiri. In Stephen R Anderson & Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 308–344. Holt, Rinehart & Winston.

Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20*, 111–176. MIT Press.

Halle, Morris & Alec Marantz. 2008. Clarifying "blur": Paradigms, defaults, and inflectional classes. In Asaf Bachrach & Andrew Nevins (eds.), *Inflectional identity*, 55–72. Oxford University Press.

Harbour, Daniel. 2007. *Morphosemantic number: From Kiowa noun classes to UG number features*, vol. 69. Springer. https://doi.org/10.1007/978-1-4020-5038-1.

Harbour, Daniel. 2008. Mass, non-singularity, and augmentation. *MIT Working Papers in Linguistics* 49. 239–266.

Harbour, Daniel. 2011a. Descriptive and explanatory markedness. *Morphology* 21(2). 223–245. https://doi.org/10.1007/s11525-010-9167-0.

Harbour, Daniel. 2011b. Valence and atomic number. *Linguistic Inquiry* 42(4). 561–594. https://doi.org/10.1162/LING_a_00061.

Harbour, Daniel. 2013. "Not Plus" isn't "Not There": Bivalence in person, number, and gender. In Ora Matushansky & Alec Marantz (eds.), *Distributed morphology today: Morphemes for Morris Halle*, 35–150. The MIT Press Cambridge, MA. https://doi.org/10.7551/mitpress/9780262019675.003.0008.

Harbour, Daniel. 2014. Paucity, abundance, and the theory of number. *Language* 185–229. https://doi.org/10.1353/lan.2014.0003.

Harbour, Daniel. 2016. *Impossible persons*, vol. 74. MIT Press. https://doi.org/10.7551/mitpress/9780262034739.001.0001.

Harbour, Daniel. 2020. Frankenduals: Their typology, structure, and significance. *Language* 96(1). 60–93. https://doi.org/10.1353/lan.2020.0002.

Harley, Heidi. 1994. Hug a tree: Deriving the morphosyntactic feature hierarchy. *MIT Working Papers in Linguistics* 21. 289–320.

Harley, Heidi & Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78(3). 482–526. https://doi.org/10.1353/lan.2002.0158.

Hawkins, Roger & Cecilia Yuet-hung Chan. 1997. The partial availability of universal grammar in second language acquisition: The 'failed functional features hypothesis'. *Second Language Research* 13(3). 187–226. https://doi.org/10.1191/026765897671476153.

Jakobson, Roman. 1958. Morfologičeskie nabljudenija nad slavjanskim skloneniem [Morphological observations on Slavic declension]. In *Selected writings II*, 154–183.

Kimball, Geoffrey D. 1985. *A descriptive grammar of Koasati*: Tulane dissertation.
Koo, Hahn. 2007. *Change in the adult phonological processing system by learning non-adjacent phonotactic constraints from brief experience: An experimental and computational study*: UIUC dissertation.
Kramer, Ruth. 2016. Syncretism in Paradigm Function Morphology and Distributed Morphology. In Daniel Siddiqi & Heidi Harley (eds.), *Morphological metatheory*, 95–120. John Benjamins. https://doi.org/10.1075/la.229.04kra.
Linzen, Tal & Gillian Gallagher. 2014. The timecourse of generalization in phonotactic learning. In *Proceedings of the Annual Meetings on Phonology*, vol. 1. https://doi.org/10.3765/amp.v1i1.18.
Maldonado, Mora & Jennifer Culbertson. 2020. Person of Interest: Experimental investigations into the learnability of person systems. *Linguistic Inquiry* 1–71. https://doi.org/10.1162/ling_a_00406.
Maldonado, Mora, Carmen Saldana & Jennifer Culbertson. 2020. Learning biases in person-number linearization. In *Proceedings of NELS 50, volume 2*. 163–176.
McGinnis, Martha. 2005. On markedness asymmetries in person and number. *Language* 81(3). 699–718. https://doi.org/10.1353/lan.2005.0141.
Moreton, Elliott. 2008. Analytic bias and phonological typology. *Phonology* 25(1). 83–127. https://doi.org/10.1017/S0952675708001413.
Moreton, Elliott & Joe Pater. 2012a. Structure and substance in artificial-phonology learning, part i: Structure. *Language and Linguistics Compass* 6(11). 686–701. https://doi.org/10.1002/lnc3.363.
Moreton, Elliott & Joe Pater. 2012b. Structure and substance in artificial-phonology learning, part ii: Substance. *Language and Linguistics Compass* 6(11). 702–718. https://doi.org/10.1002/lnc3.366.
Müller, Gereon. 2004. On decomposing inflection class features: syncretism in Russian noun inflection. In Gereon Müller, Lutz Gunkel & Gisela Zifonun (eds.), *Explorations in nominal inflection*, 189–227. Walter de Gruyter. https://doi.org/10.1515/9783110197501.189.
Müller, Gereon. 2005. Syncretism and iconicity in Icelandic noun declensions: A Distributed Morphology approach. In *Yearbook of Morphology 2004*, 229–271. Springer. https://doi.org/10.1007/1-4020-2900-4_8.
Nevins, Andrew. 2011. Marked targets versus marked triggers and impoverishment of the dual. *Linguistic Inquiry* 42(3). 413–444. https://doi.org/10.1162/LING_a_00052.
Nevins, Andrew & Paul Marty. Under revision. Expressions of paucity: Where is the upper-bound?
Noyer, Robert Rolf. 1992. *Features, positions and affixes in autonomous morphological structure*: MIT dissertation.
Payne, Doris Lander. 1985. *Aspects of the grammar of Yagua: A typological perspective*: UCLA dissertation.
Peirce, Jonathan, Jeremy R Gray, Sol Simpson, Michael MacAskill, Richard Höchenberger, Hiroyuki Sogo, Erik Kastman & Jonas Kristoffer Lindeløv. 2019. PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods* 51(1). 195–203. https://doi.org/10.3758/s13428-018-01193-y.
Pertsova, Katya. 2007. *Learning form-meaning mappings in presence of homonymy: A linguistically motivated model of learning inflection*: UCLA dissertation.
Pertsova, Katya. 2011. Grounding systematic syncretism in learning. *Linguistic Inquiry* 42(2). 225–266. https://doi.org/10.1162/ling_a_00041.

Pylkkänen, Liina, Andrew Stringfellow & Alec Marantz. 2002. Neuromagnetic evidence for the timing of lexical activation: An MEG component sensitive to phonotactic probability but not to neighborhood density. *Brain and Language* 81(1-3). 666–678. https://doi.org/10.1006/brln.2001.2555.

R Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org.

Saffran, Jenny R & Erik D Thiessen. 2003. Pattern induction by infant language learners. *Developmental Psychology* 39(3). 484. https://doi.org/10.1037/0012-1649.39.3.484.

Sauerland, Uli. 2008. On the semantic markedness of φ-features. In Daniel Harbour, David Adger & Susan Béjar (eds.), *Phi Theory: Phi-Features Across Modules and Interfaces*, 57–82. Oxford University Press.

Silverstein, Michael. 1986. Hierarchy of features and ergativity. In Pieter Muysken & Henk van Riemsdijk (eds.), *Grammatical categories in Australian languages*, 163–232. Foris Publications.

Smith, Peter W, Beata Moskal, Ting Xu, Jungmin Kang & Jonathan David Bobaljik. 2019. Case and number suppletion in pronouns. *Natural Language & Linguistic Theory* 37(3). 1029–1101. https://doi.org/10.1007/s11049-018-9425-0.

Starke, Michal. 2010. Nanosyntax: A short primer to a new approach to language. *Nordlyd* 36(1). 1–6. https://doi.org/https://doi.org/10.7557/12.213.

Starke, Michal. 2018. Complex left branches, spellout, and prefixes. In Lena Baunaz, Liliane Haegeman, Karen De Clercq & Eric Lander (eds.), *Exploring nanosyntax*, 239–249. Oxford University Press. https://doi.org/10.1093/oso/9780190876746.003.0009.

Starke, Michal. 2020. UM - Universal Morphology, UQAM: Keynote talk, North East Linguistics Society (NELS) 51.

Tabullo, Angel, Mariana Arismendi, Alejandro Wainselboim, Gerardo Primero, Sergio Vernis, Enrique Segura, Silvano Zanutto & Alberto Yorio. 2012. On the learnability of frequent and infrequent word orders: An artificial language learning study. *Quarterly Journal of Experimental Psychology* 65(9). 1848–1863. https://doi.org/10.1080/17470218.2012.677848.

Tsimpli, Ianthi Maria & Maria Dimitrakopoulou. 2007. The interpretability hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research* 23(2). 215–242. https://doi.org/10.1177/0267658307076546.

Vitevitch, Michael S & Paul A Luce. 1998. When words compete: Levels of processing in perception of spoken words. *Psychological Science* 9(4). 325–329. https://doi.org/10.1111/1467-9280.00064.

Vitevitch, Michael S & Paul A Luce. 1999. Probabilistic phonotactics and neighborhood activation in spoken word recognition. *Journal of Memory and Language* 40(3). 374–408. https://doi.org/10.1006/jmla.1998.2618.

Werker, Janet F. 1994. Cross-language speech perception: Development change does not involve loss. In Judith C. Goodman & Howard C. Nusbaum (eds.), *The development of speech perception: the transition from speech sounds to spoken words*, 93–120. MIT Press. https://doi.org/10.7551/mitpress/2387.003.0009.

Werker, Janet F & Richard C Tees. 1984. Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development* 7(1). 49–63. https://doi.org/10.1016/s0163-6383(84)80022-3.