

Learning (im)possible number syncretisms: investigating innate featural representations

Naomi Lee (New York University)

NELS 51, 6-8 November 2020, UQAM

Introduction. Feature-based theories of morphosyntactic categories like person or number seek to propose universal inventories of pieces that generate exactly the values attested across the world’s languages. They predict that natural classes characterized by those primitives should pattern together morphologically, including in which values may be neutralized to a syncretic form. Two such theories of number, Harbour (2014, 2016) and Smith et al. (2019), agree that the dual and plural values form such a natural class, but make opposing predictions for the natural class membership of the singular value. Typological data evinces the predicted dual-plural syncretism in many languages, but attestations of the other logically possible syncretisms are sparse or absent (Corbett 2000). New research, building on established findings that artificial language learners rely on natural classes (Cristià & Seidl 2008, Finley & Badecker 2009, i.a.), suggests that experimental data on the learnability of feature-based partitions of morphological paradigms can enrich our empirical landscape beyond sparse or potentially confounded typological data (Maldonado et al. 2019, 2020).

In this project, I investigate which of two proposed number feature representations’ predictions about (im)possible syncretisms better explains the performance of artificial language learners. Adults are presented with one of three “grammars”, all of which display a three-way distinction between singular, dual, and plural in the nominal domain, but show contextual syncretism in the verbal agreement for number between two of these three values. Participants are presented with a referent-selection task based only on the verbal number agreement marker (the unambiguously number-marked noun is obscured by noise in the audio stimuli). Both theories expect an advantage for the dual-plural (ABB) syncretism, but each makes a distinct prediction for which grammar should have the greatest disadvantage.

The number space: feature representation proposals. I focus on two theories of number that make opposing predictions: Harbour (2014, 2016)’s binary, cross-classifying $[\pm\text{atomic}]$ and $[\pm\text{minimal}]$ features and Smith et al. (2019)’s linear containment hierarchy. These systems both generate the typologically common set of singular, plural, and dual number values. However, these theories’ distinct notions of the sense in which features can be “shared” generate distinct natural classes within this three-way contrast.

Harbour (2014, 2016) composes number values like “singular” from bundles of binarily $[+]$ or $[-]$ -valued features. If different bundles share a valued feature, such as singular and dual both being $[+\text{minimal}]$ (i.e. having no similar subelements), they form a natural class. Even negative values of a given feature - in this case, $[-\text{atomic}]$ - can be shared by number values and therefore characterize a natural class - namely, dual and plural. Only singular and plural do not form a natural class. Although privative feature systems differ in organization from those like Harbour’s, the leading such proposal (Harley & Ritter 2002’s feature geometry) composes dual from the features characterizing singular (Minimal) and plural (Group), and so makes identical predictions: only singular-plural syncretisms are disallowed.

<i>binary, cross-classifying</i>			<i>feature-geometric</i>		<i>privative, linear</i>		
	$[+\text{atomic}]$	$[-\text{atomic}]$	INDIVIDUATION				
$[+\text{minimal}]$	singular $[+a +m]$	dual $[-a +m]$	Group <u>Minimal</u>		SINGULAR PLURAL DUAL		
$[-\text{minimal}]$	$[+a -m]$	plural $[-a -m]$					


Smith et al. (2019), on the other hand, structure the features that build up singular, plural, and dual values in a linear containment hierarchy. By the Superset Principle of Nanosyntax, subtrees (with phonological pronunciations) can be inserted if they contain the tree up to that point (Starke 2010, Caha 2019). For example, a subtree containing the PLURAL and DUAL nodes in its structure could target either “plural” or “dual” meanings. This means the structures formed by contiguous spans of adjacent features form natural classes, and that “nonadjacent syncretisms are excluded in

principle...by the *ABA theorem” (Baunaz et al. 2018:x). Thus a singular and dual syncretism to the exclusion of plural is disallowed in this kind of system.

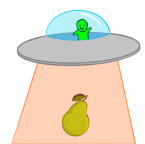
Following Maldonado & Culbertson (2020)’s proof of concept that adults retain the ability to use innate person-feature-based natural classes in learning of artificial languages, I hypothesize that syncretisms targeting shared number features should be easier to learn than those with no overlap. Both theories predict that participants learning the dual-plural (ABB) syncretism will fare well. Crucially, however, Harbour’s system (and Harley & Ritter’s) predicts a disadvantage for learners exposed to the singular-plural (ABA) syncretism grammar, whereas Smith et al.’s proposed hierarchy predicts one for learners of the singular-dual (AAB) syncretism.

Experimental methods. To test these predictions, I used an artificial language ‘ease of learning’ paradigm (Culbertson et al. 2017). Participants were taught a mini language with 6 nominal stems (for different kinds of fruit), and 2 verbal stems (meaning ‘beam up’ or ‘throw away’). Noun suffixes explicitly marked a three-way contrast for number (sg, du, pl), to promote activity of the non-native “dual” value (and all its component features) in the participants’ learning process. 2 verbal suffixes (represented as A and B) expressed agreement with the number of the object. The language had three possible grammar conditions, differing on which number distinctions were contextually neutralized (rather than systematically homophonous) in those verbal suffixes: dual-plural (ABB), singular-dual (AAB), and singular-plural (ABA) syncretism.

In the experiment, participants met friendly visiting aliens who love trying local fruits and were asked to learn their language (to help make a smoothie). Participants were first trained on nouns and their number endings, with pictures of one, two, or many of the fruit accompanying the audio as in (1). They then were trained on imperative sentences, with verbs bearing agreement markers that tracked the number of their object, with animations of one, two, or many of a fruit being beamed up or thrown away as in (2). For both the nominal and verbal training phases, training consisted of repeating audio stimuli, then making choices between two quantities of the same fruit on the basis of further audio, with feedback on their choices.

(1) seeg-po
pear-DU
 pears (dual)

(2) bice-te (seeg-cha)
beam.up-SG.OBJ pear-SG
beam up the pear!

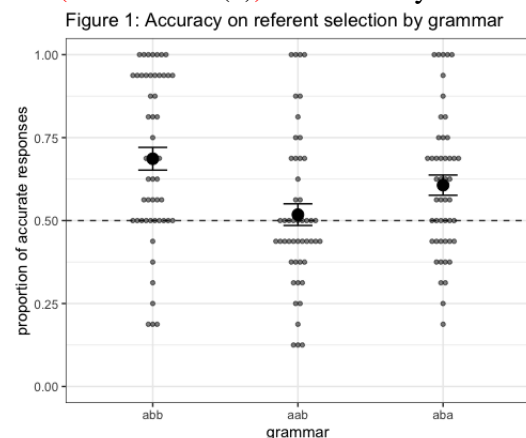


To validate participants’ success in learning the three-way nominal distinction, they were asked to choose between two quantities of fruit based on noun forms with number endings.

In the crucial test phase, to measure learning of different syncretisms, participants were again asked to choose between two quantities of fruit, but based only on imperative verb forms bearing object number agreement, with the fully disambiguating noun (bracketed in (2)) obscured by noise.

Results. 149 English-speaking adult participants, recruited from Amazon Mechanical Turk and assigned randomly between the three grammars, completed the task (including passing rudimentary sound and attention checks) (52 ABB, 49 AAB, 48 ABA). A logit mixed-effects model of responses to the noun-based validation items confirms that participants across all three grammars learned the three-way distinction (intercept: $p < 0.001$), with no significant difference between grammars.

Figure 1 shows individuals’ proportion of accurate responses on crucial test items - referent selection based on the syncretic verbal forms - by condition. A logit mixed-effects model shows that participants learning the dual-plural (ABB) syncretism gave correct responses significantly above chance (intercept: $p < 0.001$), and that accuracy in the singular-dual (AAB) syncretism was significantly lower than that baseline ($p < 0.001$). However, accuracy in the singular-plural (ABA) condition was not significantly different from the baseline ($p = 0.055$).



Discussion. Participants' similar, significant success in learning the three-way nominal paradigm confirms the learnability of non-native contrasts like the dual. Moreover, that general success with the dual, but significantly different performance on different number agreement syncretisms, corroborates the forming consensus that adult behavior in artificial language learning tasks engages innate, abstract feature representations. Namely, since dual and plural are a natural class in both theories, learners' significantly better performance on the dual-plural (ABB) syncretism supports the general relevance of feature-derived natural classes in morphological learning. Additionally, the relative disadvantage observed for the singular-dual (AAB) syncretism suggest specific support for a Smith et al.-like system where singular and dual are not a natural class.

Selected references:

- Harbour, Daniel. 2014. "Paucity, abundance, and the theory of number." *Language* 90.1: 185-229.
- Harbour, Daniel. 2016. *Impossible persons*. MIT Press.
- Harley, Heidi & Elizabeth Ritter. 2002. "Person and Number in Pronouns: A feature-geometric analysis." *Language* 78.3: 482-526.
- Maldonado, Mora & Jennifer Culbertson. 2019. "Something about us: Learning first person pronoun systems." In A. C. Goel, C. M. Seifart, C. Freksa (eds.) *Proceedings of the 41st Annual Meeting of the Cognitive Science Society*: 749-755.
- Maldonado, Mora & Jennifer Culbertson. In press. "Person of interest: Experimental investigations into the learnability of person systems." *Linguistic Inquiry*: 1-71.
- Smith, Peter W., Beata Moskal, Ting Xu, Jungmin Kang, & Jonathan David Bobaljik. "Case and number suppletion in pronouns." *Natural Language & Linguistic Theory* 37.3: 1029-1101.