The Acquisition of Western Armenian Number Features with Variable Input

1. Introduction

In Western Armenian, plural marking on the noun is optional for numerals greater than one, with certain connotative differences depending on semantic context, and the optionality of plural marking is affected by the presence or absence of complementizers, classifiers, and the intricacies of count-mass distinctions.

In many different syntactic environments, so-called singular nouns in Western Armenian have an inclusive semantic denotation (Bale & Khanjian, 2014), and thus resemble bare nouns that are assigned a general number interpretation in other languages (Corbett, 2000). However, following Martí (2020)'s account, unlike other languages whose NPs are either strictly [±atomic] (e.g. English, French, Spanish) or strictly [±minimal] (e.g. Hungarian, Finnish, Turkish), Western Armenian, along with a handful of other studied languages such as Itzaj Maya and Miya (Martí 2020), can be [±atomic] or [±minimal] (but the precise scope of this optionality and its acquisition have not been explored, neither for children nor adults). It follows that in certain limited grammatical contexts the bare nouns do sometimes have a strict singular meaning, and the plural is semantically very marked; other complicating factors are that unlike in English or French, "when a plural noun appears in an *if*-clause, in a *yes-no* question, or in the restrictor of a universal quantifier, the noun is not able to quantify over singular individuals" (Bale, Gagnon & Khanjian 2011).

In this paper, after a literature review and through a series of child experiments (subdivided into age cohorts) on context-based production and comprehension, we explore how the semantic properties of Western Armenian number features are acquired and more specifically, uncover the precise scope of where [±atomic] or [±minimal] traits appear in nouns and the reasons triggering such interpretations at the various stages of language acquisition. Our hypothesis is that, roughly following the proposed order of acquisition of features in Hanson, Harley & Ritter (2000), Matsumoto (1985), Austin (2013), Bel & Rosado (2009), Western Armenian-speaking children ought to produce bare nouns before make use of atomicity and minimality via plural marking and classifier use.

2. Explanation of Phenomenon and Data

Western Armenian plurals have a strict plural interpretation, while singulars have a general number interpretation, and this may be a consequence of the plural feature being marked both semantically and morphologically (but there are sometimes verbal number mismatches which are permitted). Depending on context, the use of the plural marker for a noun already overtly marked with a numeral can sound odd:

(1)	Šiš unim.	bottle-SG have-1SG	'I have a bottle/bottles'
(2)	Yergu šiš unim	two bottle-SG have-1SG	'I have two bottles'
(3)	Yergu had šiš unim	two CL bottle-SG have-1SG	'I have two bottles'
(4)	Šišer unim.	bottle-PL have-1SG	'I have bottles.'
(5)	#Yergu šišer unim	two bottle-PL have-1SG	'I have two bottles'
(6)	??Yergu had šišer unim	two CL bottle-PL have-1SG	'I have two bottles'

Note that the number 'zero' is unavailable in Western Armenian in prenominal position, which is rare cross-linguistically (Bylinina & Nouwen, 2018). For mass nouns, the scope of the plural is very restricted. Furthermore, plurals do not behave the same way as English plurals in downward-entaining

contexts, which is a large topic that we will only briefly touch on. A starting point for our analysis will be the set of denotations below (adapted from Bale & Khanjian 2008):

[[**šiš**]] = {x: x is a bottle or a group of bottles}
[[**šiš**$_{\textbf{count}}$]] = {x: x is a bottle}
[[-**er**]] = λP {X: X ⊆ P & |X| >1}
[[(**šiš**$_{\textbf{count}}$) –**er**]]= {X: X ⊆ [[**šiš**$_{\textbf{count}}$]] & |X| > 1}

The plot thickens when we add conjunctions – for 'John and Arpi are teachers', we have four possibilities, only two of which are felicitous ((7) and (8)), (9) is only valid with a bizarre symbolic meaning of both of them acting as one teacher in unison, and we get an exclusively plural interpretation for sentences like (11), where one would not stand up if they had but one child:

(7) Hovannes-ə yev Arpi-n usutsič en.
John-DEF-DET and Arpi-DEF-DET teacher are.

(8) Hovannes-ə yev Arpi-n usutsič-ner en.
John-DEF-DET and Arpi-DEF-DET teachers are.

(9) #Hovannes-ə yev Arpi-n meg usutsič en/e.
John-DEF-DET and Arpi-DEF-DET one teacher are/is.

(10) *Hovannes-ə yev Arpi-n usutsič mən en/e.
John-DEF-DET and Arpi-DEF-DET teacher-IND-DET are/is.

(11) Amen mart vor bzdig-ner une-r vodk-i gayne-ts-av/-an.
all person that child-PL have-PAST-3SG foot-DAT stand.up-PAST-3SG/3PL.
'Everyone who had (two or more) children stood up.'

Sentence (7) and (8) are in free variation and can be selected for stylistic effect; in contrast, it is not possible to get a coherent interpretation of (10), even though such an interpretation should be available if a distributive operator were present ($^{\text{DIST}}P = *\ ATOM(P)$, where *ATOM(P)* is the subset of *P* that includes only atoms, i.e. an operator that distributes the NP-predicate over each member of the plural-subject) (Landman, 2000). Note that when appearing in predicate position, a noun such as *usutsič* 'teacher' can serve as a predicate to both singular subjects (say, *Hovannes-ə usutsič e*, 'John is a teacher'), plural subjects, or plural subjects such as conjoined NPs such as:

(12) Dəʁa-ner-ə yev aʁčig-ner-ə ašagerd en.
boy.PL.DEF-DET and girl.PL.DEF-DET student are.
'The boys and the girls are students.'

Sentence (12) stands in sharp contrast with English, where "the bare noun can only be predicated coherently of singular objects" (Bale, Gagnon & Khanjian 2011), hence why a sentence like **John and Arpi are a teacher* cannot be felicitous.

Another complicating factor is that, contrary to the strong cross-linguistic generalization that classifier languages (Mandarin, Japanese, Korean, etc.) lack systematic number marking and that number marking languages (English, French, German, etc.) lack classifiers, all varieties of Armenian

have both a classifier system and a number marking system. For example, as we can see below, adding a classifier will block any possibility of adding a postposed determiner or plural marker:

(13) Madid-mə knetsi.
pencil-INDEF-SG buy-PAST-1SG
'I bought a pencil.'

(14) Yergu madid / #madidner knetsi.
two pencil-SG/-PL buy-PAST-1SG
'I bought two pencils.'

(15) Yergu had madid (knetsi)
Two CL pencil-SG buy-PAST-1SG
'(I bought) two pencils.'

(16) *Yergu had madid-ner
Two CL pencil-PL
'two pencils'

(17) *Yergu had madid-mə
Two CL pencil- INDEF-SG
'two pencils'

(18) Yergu madid-ner-ə knetsi.
Two pencil-PL-DEF-DET buy-PAST-1SG
'I bought the two pencils'[1]

Thus, although Western Armenian has both number marking and classifiers, both cannot appear within the same NP construction (Bale & Khanjian 2008), in contrast to Nez Perce, whose numeral suffixes do not behave like those of Western Armenian in disappearing when plural is present; instead, Nez Perce NPs with classifiers consistently take the plural forms when combining with numerals higher than one (Deal, 2013). Thus, this is strong proof that morphological plural markings and the presence of classifiers do not have to be in complementary distribution cross-linguistically.

3. Brief Literature Review and Theoretical Considerations

We can roughly categorize the literature into three eras – the early era (1960s – mid-1990s), where there was nothing yet on Western Armenian, but a lot of theoretical work on internal DP structure (Ritter 1991) which would later become useful. The second era, approximately from the late 1990s to 2008, sees some authors posit that perhaps Western Armenian is just like Turkish, such as Sigler (1997), Ackermann et al. (1998), and Lonsdale & Danielyan (2004). In the third era (2008 – today), authors are essentially claiming that this language is doing its own thing (mostly based on the work of Bale, Gagnon, and Khanjian (2008, 2010, 2011, 2013, 2014)), with only certain elements superficially mimicking the Turkish pattern.

1 Added this example after Chaitanya Malaviya's feedback – this sentence would typically refer to a specific pair of pencils, hence the determiner use, and the can optionally be pluralized morphologically as shown in this example.

Hrayr Khanjian (MIT), Michael Gagnon (Maryland), and Alan Bale (Concordia) are the three linguists responsible for first delving into plural marking and the semantics of number interpretations in Western Armenian in the mid-2000s. Khanjian & Bale (2008) discuss Western Armenian classifiers and number marking in a handout for a *SALT* proceeding and give predictions as to the correct interpretation of sentences like in (11).

By 2011, Khanjian and Bale gave their empirical and theoretical accounts of grammatical number in the numeral+noun constructions in Turkish, Western Armenian, and English, which is modified (and according to Martí (2017), "substantially improved") by Scontras (2013). Martí (2020) then proposes a Scontras-based system that accounts for the data with a uniform semantics for numerals, an empirically justified semantics for Turkish and Western Armenian nouns, and an adequate understanding of exclusive and inclusive plurality within Harbour's theory (a theory which we will not get into, but suffice to say that he proposes a theory of grammatical number where the source of sensitivity to minimal parts is the feature [±minimal] as distinct from [±atomic]).

Scontras also included measure terms in his analysis (like "one *kilo* of flour"), but ignored classifiers altogether, which highlighted "the breadth required by the semantic mechanism that modulates grammatical number" (Scontras, 2014). His syntactic analysis included a secondary projection of the head of NumP, in-between which either SG or PL would be inserted, as an operator that "establishes conditions on the denotation of the resulting nominal: SG checks for singularity of the predicate, and PL applies when singularity is not satisfied" (*ibid*.). Scontras summarizes his findings that zooming in on the measure feature μ in his proposed semantic denotation of plural marking:

[[**SG**]] = λP: $\forall x \in P[\ \mu(x) = 1\ ].\ P$
[[**PL**]] = λP. P

where $\mu_{P\text{-}atom}(y)$ is defined only if $y \in P$; when defined: $\mu_{P\text{-}atom}(y) = |\{x \in P: x \leq y \ \& \ \neg \exists z \in P[z < x]\}|$[2]

Moreover, Wolf (2013) uses the Optimality Theory framework to explain number marking on possessed nouns in Western Armenian, but in order for his derivations to function as intended, he posits "a phonologically null singular affix, as opposed to saying that singulars carry no number morph at all" (Wolf 2013:154).

Theoretical basis of number marking features

As explained by Scontras (2014), speakers of number marking languages decide between singular and plural forms of nouns as they embed them in larger linguistic contexts, but how a language's number marking features interplay with numerals, quantifiers, determiners, classifiers, and the morphology can differ vastly cross-linguistically.

2 "Cross-linguistic variation in patterns of number marking falls out once we allow variation in the selection of μ: in English μ is relativized to the head of #'s sister; in Turkish, where all numerals occur with singular-marked nouns, μ is relativized to the phrasal complement of # on the basis of P-atoms. In Western Armenian, where the pattern of number marking is intermediate between the English and the Turkish systems, there is optionality in the selection of μ: either the head or the phrasal approach may apply. Our account of this variation makes do with a uniform syntax and semantics for numerals across these languages (cf. the variation in numeral semantics proposed in Bale et al., 2011a) within a standard semantics framework (cf. the OT account of Farkas and de Swart, 2010)." (Scontras, 2014, p. 560-561)

Bale, Gagnon, and Khanjian (2011) argue for an account of the three feature patterns mentioned in the Introduction, where both the semantics of numerals and the semantics of nouns may vary from one language type to another: numerals in different languages may have either subsective or intersective[3] semantics, and Martí (2020) notes that the numeral *one* may or may not have the same kind of denotation as other numerals within the same language, and morphologically singular noun may also have a different semantics in different languages (singular in English number, neutral in Turkish or Armenian-type languages). Martí (2020) addresses the empirical and theoretical shortcomings of Bale, Gagnon, and Khanjian's account by focusing on Armenian possessive NP constructions (among other languages), and combining their findings with those of Harbour and Scontras to come up with a simplified system in which numerals greater than *one* combine with morphologically plural nouns in English-type languages ([±atomic]) because the members of a set of non-atoms are not atoms, and such a set can thus be characterized as [-atomic], which spells out as *-s* in English, whereas such numerals combine with morphologically singular nouns in Turkish because individuals in set of non-atoms also count as having no subparts in the set, and such a set can thus be characterized as [+minimal], which spells out as a null morpheme in Turkish-type languages. For Armenian-type languages, the number system has both [±atomic] or [±minimal] at its disposal, though these features are activated differently for different NPs according to Martí (2020). The proposed trees below summarize these findings, while also allowing for optionality.

Furthermore, the number marking features in Western Armenian to be in grammatical competition, resulting in a restricted interpretation of singular nouns rather than plurals, though this type of competition only occurs in certain grammatical environments: "namely, when a noun phrase has a definite interpretation but critically not when it has an indefinite interpretation" (Bale & Khanjian, 2014). It remains an open question if syntactic complexity *per se* would change the felicitousness of certain interpretations.

As we can see, the literature on this issue is quite recent and fraught with disagreement (especially Scontras and Martí), especially on more complex issues like quantification in downward entailment contexts. Bale and Khanjian have repeatedly revisited this and related topics over the past decade. We also do not have any acquisition studies here. One particular weakness is that we do not have monolingual Western Armenian speakers (this assumption can be taken to be true, see Donabedian-Demopoulos 2018 for a historical and sociolinguistic sketch of the diaspora), so the data provided by speakers always has the risk of being tainted by other languages they are co-native or fluent in.

Another important assumption here is that the morphological diagnostic which applies to Western Armenian is that of null marking (Bale, Gagnon & Khanjian 2010). As is clear from sentences (7), (8), and (11), the plural feature is associated with overt phonological content (*-er* for monosyllabic words, *-ner* for all polysyllabic words, in a predictable allomorphic relationship) whereas the singular feature is not, suggesting that singular is morphologically unmarked whereas plural is marked even semantically.

A small note regarding the dominant dialect of Armenian – Eastern Armenian, which has had different superstratal influences (as it belongs to the Caucasian Sprachbund, and not the

3 As explained by Bale, Gagnon, and Khanjian (2011, p. 585): "[t]he adjective *pregnant* is a prototypical intersective modifier. To determine the interpretation of the NP pregnant mouse, it is sufficient to intersect the set of all the things that are pregnant with the set of all the mice. In other words, pregnant can be interpreted as a set and restrictive modification can be achieved through intersection." Subsective modifiers have a relative interpretation and cannot be straightforwardly assigned a set interpretation.

Ottoman/Balkan one, and has had more influence from Russian and far less from Turkish) and many of the generalizations made about Western Armenian cannot be applied to Eastern. For example, Paperno (2012) gives us this sentence:

(19) Ov ew owm ēr eʁan-ov xp'owm
who.NOM and whom.ACC AUX pitchfork.SG.INSTR beat.PRES
'Who struck whom with a pitchfork?'

And Paperno notes that his Western Armenian informant failed to approve this kind of hybrid coordination and did not share the same interpretation as the Eastern speakers. Unfortunately, there is precious little research on the semantics of number marking in Eastern Armenian, but the few speakers known to the author have claimed that they would take issue with some of the constructions used in this paper that Western speakers find acceptable.

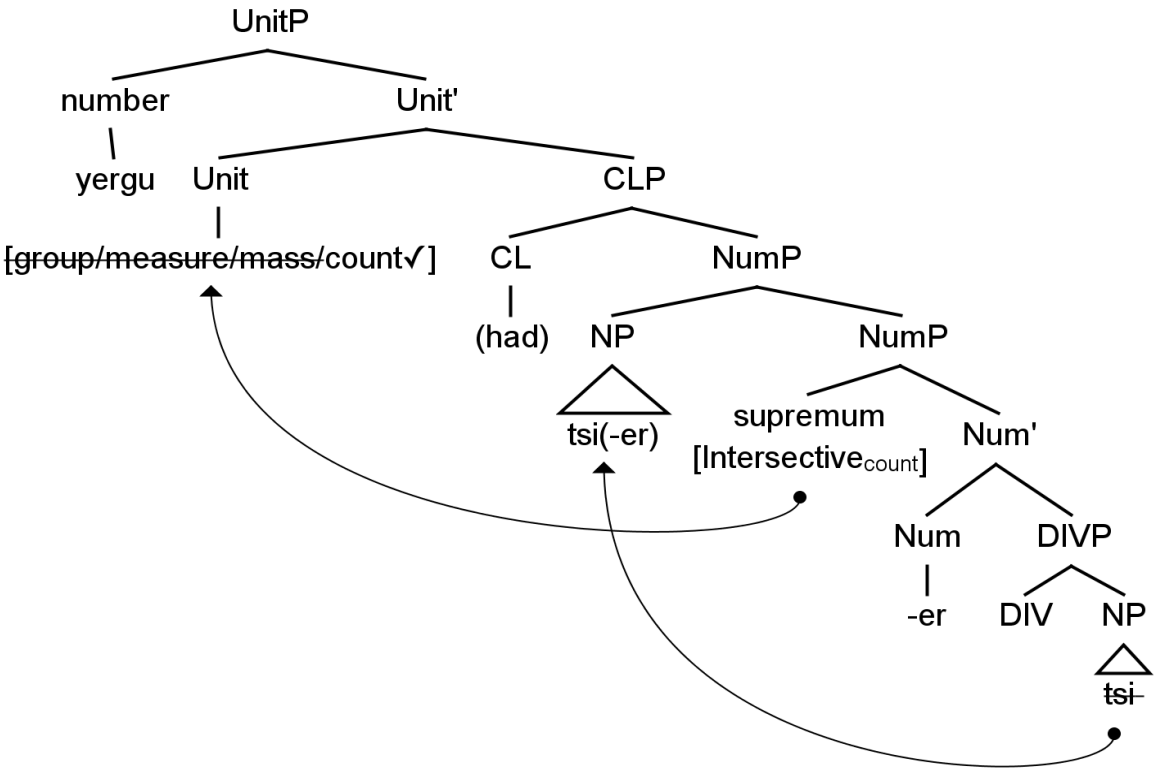

Figure 1 – Tree structure, arranged (and slightly modified) by summarizing the various findings of the literature, showing us the grammatical optionality available in *yergu (had) tsi(-er)*, 'two horses'.

The presence of classifiers complicates the matter for Western Armenian, though the classifier is added to the tree structure for the sake of completeness. Let us take the word [[**tsi**]] (horse) – in the proposed tree structure, at the point where the supremum (which bears an [intersective$_{count}$] feature) combines with the NP, the two structures induce different presuppositions. For the supremum of [[**tsi**]] to be defined, there must be at least one horse in the denotation of [[**tsi**]] – in other words, the function induces the presupposition (among others) that at least one horse exists. In contrast, for the supremum of [[**tsi-er**]] to be defined, there must be a group consisting of two or more horses in the denotation of [[**tsi**]]. Put in other words, the function induces the presupposition (among others) that at least two horses exist in Figure 1, but the number of horses in the context of Figure 2 below remains undefined.

While the contexts in which number-neutral nominals (like English '*do you have children?*') can appear in Western Armenian cannot be defined in purely semantic terms, such nominals are clearly syntactically selected according to Pereltsvaig's (2013) account.

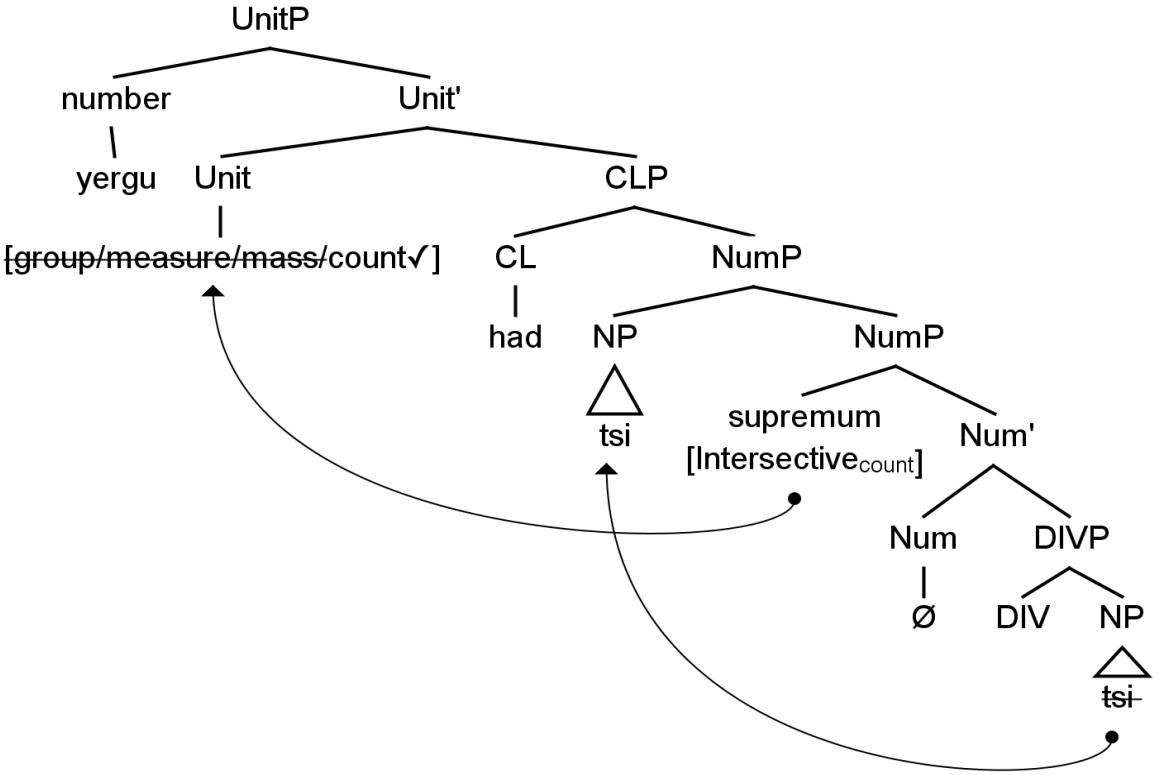

Figure 2 – Bare plural tree structure

Thus a second NumP node would allow for some functional element to assign range to more than one value (following Potts (2008) on fitting classifiers and plural marking), thus respecting optionality and allowing for both a singular and plural surface realization. Whether in this language DIV (Division) ought to be bundled with CL (as in Korean[4]) is an issue for another paper.

Thus, following Bale & Khanjian (2014), in cases where [[**yergu**]]([**[tsi-er**]]) contains a supremum ("σ"), σ[[**yergu**]]([**[tsi-er**]]), the modifier *yergu* restricts the singleton set to groups with a cardinality of two. If the supremum contained within [**[tsi-er**]] is a group with only two members, then the result of this restriction will not be empty and the result of applying the function will be the two-member supremum. If the supremum contained within [**[tsi-er**]] is not a two-member group, then the result of the restriction will be empty and hence the function will not be applied, and its interpretation will be undefined. The presence of a classifier will block the addition of any over plural (or definite) morphological suffix to the word (for most speakers), thus will receive either a singular interpretation

4 "In Korean, the demonstrative morpheme clearly applies after numeral modification. Like Western Armenian, Korean has a plural marker. It also has singular nouns with a general number interpretation as well as plural nouns with a strict plural meaning. Critically, singular nouns in Korean have a strict singular interpretation within the context of a definite DP (actually a demonstrative DP since Korean does not have nondemonstrative definite markers). However, unlike in Western Armenian, in Korean singular nouns can combine with numerals within definite DPs (i.e., demonstrative DPs)." (Bale & Khanjian 2014, p. 24.)

or a bare plural interpretation, likely dependent on context (and the presence or absence of a numeral), as seen in Figure 2.

Thus, these proposed tree structures for the internal structure of DPs/UnitPs can at least conceptually capture the full extent of optionality in number features in Western Armenian, along with more complex feature-based interactions with count/mass (and other) distinctions and a classifier system overlay. The summarized findings from the literature also justifies why, unlike English plurals, Western Armenian bare plurals do not obligatorily scope under negation (*Bzdig-ner chi vaze-ts-in* (‘children didn’t run’) is ambiguous, meaning either there are some children who did not run (although others might have) or there are no children that ran (Bale & Khanjian 2009)) if they receive a subsective feature. Ideally, the long-term goal of acquisition research in this topic would be to find out how children acquire this DP-internal structure.

Additional Remarks

As mentioned in the Introduction, there are languages like French or English where numerals greater than *one* obligatorily combine with plural nouns, there are languages like Hungarian or Turkish where such numerals obligatorily combine with bare nouns, and finally there are languages like Western Armenian where such numerals optionally combine with either bare nouns or plural nouns (Bale, Gagnon & Khanjian 2011). Of theoretical importance here is that traditional Gricean reasoning yields the wrong empirical predictions for this language (*dəʁa vaze-ts*, boy-SG run-PAST, $\exists x.\text{RAN}(x)$ & $\text{BOY}(x)$ & $|x| \geq 1$, Bale & Khanjian, 2014), and that Western Armenian seems to go against a general trend which permits quantification over individuals as well as pluralities when plural nouns appear in downward entailing contexts. There seems to be something about having the possibility of both [±atomic] and [±minimal] traits that rearranges how number features are dealt with in the semantics; for example, in sentence (11), we can see that Western Armenian plurals do not behave in the same way as English plurals in downward entailing contexts, as its plurals should never allow for reference to singularities, since one can remain seated if they only had one child (and would indeed be expected to not have obeyed that command if the sentence were posed in the imperative mood).

Turkish-type languages (meaning [±minimal]) use a subsective semantics for any numeral greater than one, an intersective semantics for one, and a different semantics for nouns, whereas Western Armenian is like Turkish-type languages but its numerals are always intersective:

$$[[\textbf{yergu}]] = \{x : \exists\, Y\, (Y \in \text{PART}(x)\ \&\ |Y|=2\ \&\ \forall z(z \in Y \rightarrow \text{ATOM}(z)))\}$$

where $\text{ATOM}(x) = 1$ iff x is an atom in the domain of the model – iff $x \in D$ & $\neg\exists z \in D.(z<x)$

4. Proposed Set of Experiments

For both sets of experiments, the data (based on the sentences above) will be compiled with their features ([±atomic] or [±minimal]), number, morphological agreement, classifier presence, verbal agreement (if applicable) and then statistically analyzed for tendencies and age cohort-related acquisitional milestones. The data, both elicited and shown, are to be kept simple. As with research on any heritage or endangered language, special attention has to be given to the possibility that the data

provided by participants may be at risk of being tainted by other languages they are co-native or fluent in[5].

Participants

Our project seeks to engage at least 18 children as participants, broken down into three cohorts, ages 3, 4, and 5, and one adult cohort (thus, 6 adults total); preferably six for each. Adult (like parents of the child participants) participants are to be recruited from social media contacts and groups (such as Western Armenian-language elementary schools) who have expressed an interest in transmitting or preserving their heritage language in their families and from advertisements in online magazines and newspapers, all within the Washington-Boston corridor as there are a number of Western Armenian-speaking communities in those areas. All adult subjects will be compensated at a rate of $10 per hour. Both male and female adults of any age are included. Adult participants are only to be included if they are sufficiently orally fluent in the language as to be able to hold a conversation fluently, but literacy in Western Armenian (which has its own non-Latin script) is not required.

Child participants are eligible only if they have at least a modest command of Western Armenian (the self-directed questionnaire given to parents would clearly ask this). Tasks will be broken down into manageable chunks, with frequent breaks. Child participants are compensated with a small bag of stickers and corn syrup-free candies. Parents will be allowed to be present in the laboratory but will be instructed to not give answers or guide any decisions.

Production

Our methodology here would involve having simple pictures (with audio) on a computer screen involving scenarios with a cartoon character that we would introduce, let's call them Garo, holding none, one, two, or three of a small number of objects such as familiar, tangible things like pencils, apples, books, etc., and asking the cohorts of children "how many of X does Garo have?", while instructing them to answer using a full sentence (the noun, shown as X here, in the bare form at first, with later examples using the morphological plural marker). This will be done by having a few sample scenarios with their question already answered in the bare form. We recognize that the sample sentences may skew the results, which is why bare forms are preferred as they are, according to the literature, the least marked. The numbers of objects are to be kept small as to not make this task into a difficult counting task, especially for the younger cohorts (cf. Huang et al. 2013, Carey & Barner 2019).

We provide sample stimuli in Table 1; as mention by Weiqiu You, the questions' form in the production task might influence the participants' reactions, which is why counterbalancing with different types of word forms in the question (involving slightly different ways of phrasing the sentence, or using tangible, specific nouns with different physical properties) as well and see if there is an influence. Helen Jin and Aura Cruz Heredia also raise the point that eliciting full sentences for the youngest cohort may be a challenge – which is why it may be a good idea to give a few minutes' worth

5 To address Andrew Wu's concern here – one may to mitigate this as part of the experimental design is to be speaking/instructing participants in Western Armenian as opposed to English (as to prime their minds, to the extent possible, not to be thinking in English and merely translating), which is how potential trans/inter-languaging effects might be detected. Another way to detect English influence would be to repeat this experiment in areas where there are sizable Western Armenian communities, such as France or Québec, where French is the dominant language, or Lebanon, where Arabic is dominant even in minorities, as has been done for voicing and voice-onset timing (Kelly & Keshishian 2019).

of training with sample answers (using full questions), and asking the participants to replicate this pattern.

Scenario shown on screen with audio prompt	Tokens (n=40, grand total 40x24=960) using different nouns	Expected child answer	Expected adult-like answer
How many books (bare, [+atomic, -minimal]) does Garo have (when 2 or more)?[6]	10	Numeral + very rare classifier use + bare plural	Numeral + mix of bare plural and morphological plural + common optional use of classifier
How many pencils [bare, [+atomic, +minimal]] is Garo holding (when 1)?	10	Numeral or none + no determiner + singular + rare classifier use	Num + Indefinite determiner + singular + common optional use of classifier
How many books (pluralized, [-atomic, -minimal]) is Garo carrying (when 2 or more)?	10	Numeral + very rare classifier use + overt plural	Numeral + mix of bare plural and morphological plural + common optional use of classifier
How many books (pluralized, [-atomic, +minimal]) is in front of Garo have (when 1)?	10	Numeral + no determiner + singular + rare classifier use	Numeral + possible determiner + singular + common optional use of classifier

Table 1 – Sample stimuli with anticipated results.

Comprehension

In brief, this small set of comprehension tests involves making children hear variants of sentences (1)-(6), (7)-(10), and (13)-(18), while the experimenter asks them to rate the acceptability of each variant. In these sentences, there will be shown a picture for each logical possibility (scenarios containing answers with and without the morphological plural marker for the numerals 1, 2, and 3, with or without the classifier, etc.). There will be 40 tokens per participant (960 tokens total, as the production task above). A more adult-like judgment will be deemed as being comprehended better and/or sounding more felicitous. A 1-7 (7 being utterly ungrammatical) judgment scale or a scale of similar purpose[7] more suitable for younger children will be used to judge which variants are deemed to be more or less acceptable. For the younger cohort of children, we expect to see their results roughly line up with the ages of acquisition of various features examined by the literature as summarized in Table 3, whilst simultaneously exhibiting a more chaotic (random, less defined) pattern where roughly all grammatical and ungrammatical judgments do not clearly pattern with the adult judgments.

6 Languages such as English and French have [±atomic] number systems, where [+atomic] is not morphologically realized overtly, but [-atomic] is, as –s.

7 Following Lea Mangifesta, Aura Cruz Heredia, Wenting Cai, Chaitanya Malaviya, and Helen Jin's advice, a scale that uses images (such as happy vs sad/confused faces) may be more suited for the younger children than the number scale, in which case we would use an equivalency scale between the seven faces (from happiest to saddest or most confused, most distraught) as to be able to quantitatively match our two scales for statistical analysis. This would make the answers section of our comprehension task more child-friendly.

Sample sentence	Features	Anticipated results (in adults)
Garo-n madid prnets.	N	1
Garo-n madid-ner prnets.	N-PL	1
Garo-n madid-ner-ə prnets.	N-PL-DEF	1
Garo-n madid-mə prnets.	N-INDEF	1
Garo-n had madid prnets.	CL N	7
Garo-n had madid-mə prnets.	CL N-INDEF	7
Garo-n meg madid prnets.	NUM1 N	1
Garo-n meg had madid prnets.	NUM1 CL N	1
Garo-n meg had madid-ə prnets.	NUM1 CL N-DEF	3
Garo-n meg had madid-mə prnets.	NUM1 CL N-INDEF	5/1[8]
Garo-n meg madid-ner prnets.	NUM1 N-PL	7
Garo-n meg had madid-ner-ə prnets.	NUM1 CL N-PL-DEF	6/1[9]
Garo-n yergu madid prnets.	NUM2 N	1
Garo-n yergu had madid prnets.	NUM2 CL N	1
Garo-n yergu had madid-ə prnets.	NUM2 CL N-DEF	6
Garo-n yergu had madid-mə prnets.	NUM2 CL N-INDEF	6
Garo-n yergu madid-ner prnets.	NUM2 N-PL	1
Garo-n yergu had madid-ner-ə prnets.	NUM2 CL N-PL-DEF	4

Table 2 – Sample stimuli of sentences that pattern with the examples seen in (13)-(18), with the base “Garo held pencil” sentence, with adult anticipated results . The -n suffix on Garo is the obligatory definite determiner.

For simplicity’s sake, Table 2 only shows us different combinations of the same basic construction – in the data to be given to the participants, we will be mixing nouns in order to make it more interesting, and not ordering them according to any pattern (also mixing the SOV pattern), as anti-repetition and anti-boredom measures.

Features / Age Cohorts	3-4	4-5	5-6	Adults	Predictions from literature
N (=PL)	4	2	1	1	Bel & Rosado (2009:200), SG before PL
N (=SG)	1	1	1	1	Hanson et al. (2000:5) if [±minimal]
NUM1 CL N	5	3	2	1	Matsumoto (1985:82), CL mastered later
NUM2 CL N	5	3	2	1	Yamamoto (2000:385) for NUM+CL
NUM2 CL N-PL	7	6	5	4	Harley & Ritter (2002) for morphological PL

Table 3 – Summarized anticipated results broken down by age cohort

8 The expression “meg had”, depending on context and stress (bearing a single stress, as opposed to the two expected stressed syllables in a NUM + CL construction), can also mean “at once/suddenly”.

9 Same potential effect as above footnote.

Statistical methods and miscellaneous issues

In terms of statistical methods, given that we have more than one variable (features, classifier use, morphological plurality), and since the performance of each cohort cannot be assumed to be normally distributed, we aim to use Pearson's correlation test and for extra precision, we intend to run the nonparametric equivalents to our test (primarily Spearman's rank correlation coefficient and Fisher's exact testi). The data is to be extracted and elaborated in LibreOffice Calc, and the statistics software to be used is a combination of Jamovi and Jasp. The hardware we have is a Yeti Blue Microphone via USB 2.0, using a PCM S16 LE (s16l) codec, mono channel, at a sampling rate of 44100 Hz with 16 bits per sample, in a small-sized room with a carpet and minimal clutter, in a university setting. There are no other practical or logistical difficulties in obtaining the data. Exact acoustic measurements are immaterial, since all we are concerned with here are the prompted answers, either in production or the judgment tasks for comprehension, and since there can be no quantitative outliers, there is also no need to crop or clean up the data.

Regarding this project's feasibility – the timeframe is to be kept under one year (including recruitment, laboratory data production, analysis, and report) and the project can easily be converted to the online sphere if needed due to pandemic precaution measures. Moreover, due to the nature of our design, the methodology is highly reproducible for future studies (and it can easily be adapted to other mixed classifier/numeral plural system languages) and the cost estimate can be kept relatively low.

Contributions toward pursuit of long-term goals

The proposed research represents an important first step in accomplishing our long-term goals, which are: 1) to systematically study the morphological system of Western Armenian, including aspects of acquisition, contact language effects, and variation; 2) the building of digital corpora for lesser-studied and endangered languages for ease of access and more research opportunities for future linguists (with consent from anonymized participants); and, 3) for further research, other factors such as negation, tense/aspect, intradialectal variation, and sociolinguistic pressures, can be considered.

Possible problems

Other than the age-related issues already mentioned, there are several potential weaknesses to our proposal, some of which are intractable, such as the small number of competent Western Armenian speakers almost anywhere in the world, including the United States (with a minor exception of some parts of Southern California (Chahinian & Bakalian (2016)). Given that Western Armenian is a heritage language, another issue is interference or influence from other languages in which the participants are co-native. In the healthiest of immigrant speech communities, very young children living in a home where both parents speak the minority language may use it as their primary means of communication, their experiences of the dominant language (English in our case, but also Spanish for Argentina and French for France and Belgium, which are other areas Western Armenian speakers live) limited to overhearing it or being spoken to. There is also the effect of birth order – first-borns tend to be the most fluent in the heritage or immigrant language (Godson 2004a); later children often prefer to use the dominant language that their siblings had learned in school (*ibid*.). Second, third, and fourth-generation Western Armenian speakers may live in families where the native language is not actively used – in such cases, children may be passively exposed to the minority language as overhearers or by being spoken to; they may understand but never speak it (Godson 2004b). Nevertheless, Western Armenian speakers are considered to have relatively high retention rates compared to most other groups in

Lebanon (Filian 2018) and Canada (Karapetian 2014, Buda 1992), though less so in the US (Karimian 2020).

A redeeming quality of a heritage language is that it allows us to study variable input (especially in morphological forms in such languages with relatively rich verbal or nominal morphology) better than in monolingual, linguistically stable situations, as heritage language speakers tend to exhibit greater individual variability (Montrul 2016). Especially for situations involving all three elements under study, namely classifiers, numerals, and overt plurals, we expect to see a somewhat large range of adult grammaticality judgments.

Protection of human subjects plan

Risks: There are no known or anticipated risks associated with participating as a human subject in the proposed research. All sounds and images exposed to the adult and child participants are well within the norm in the human everyday listening environment; as such, there is no risk of damage to the skin, eyes, auditory system, immune system, or brain by the stimuli presented in these experiments. All participants or their caregivers are instructed that they can terminate a session at any time, for any reason, and without penalty. At the lab, child participants may be accompanied by an experimenter and/or care-givers at all times to further minimize risk. All adults must sign a consent form, and all parents or guardians need to give express permission for their children to participate in this study. When possible, the parents or guardians are debriefed after data collection. Chances of boredom or discomfort in the lab for younger participants are slight, given that we make every effort to make the lab environment appealing for most children. There is no risk of discomfort lasting beyond the testing session.

Benefits: Other than the aforementioned $10/hour compensation to the adult participants, there are no other direct benefits to the participants themselves – the chief indirect benefit is that they would be contributing to the body of knowledge in this field.

5. Conclusion

We have given an overview of the extant (mostly theoretical and descriptive) literature and proposed a set of experiments that would shed light on a crosslinguistically rare combination of number features. We have seen that languages can fall into three types of number systems under the [±atomic] and [±minimal] feature scheme, though the three types still surface as to what descriptively appear to be singular-plural systems (Martí, 2020) – a superficially simple concept can actually have a lot of nuances which show up in all sorts of interesting ways in a language's morphology, phonology, syntax, and ultimately semantics. Previous theoretical or acquisition accounts (for other languages) only dealt with classifiers separately from number marking semantic and syntactic properties, while our experiment incorporates the classifier system into the number marking system. The child acquisition results would further guide us to a better understanding of these issues. There remain plenty of venues for further theoretical research, such as how to merge the above hypothesis with the extended DP domain, possessive phrases, non-SOV word order (Western Armenian has a slight preference for SOV but can accept any word order for most types of sentences), and how Western and Eastern Armenian differ in these respects.

For future acquisitional research, it would be worthwhile to further explore downward-entailment, inclusive vs. exclusive and subsective vs. intersective interpretations; children's

comprehension of plurality under negation (and compare the results with Turkish, Renans et al. 2020[10]), and other factors such as tense/aspect, intradialectal variation, and sociolinguistic pressures. With our results, we will be able to analyze children's preference for certain forms, and that can tell us more about acquisition process of complex DPs that include numerals, classifiers, and morphological information[11].

10 Following Alexandros Kalomoiros's recommendation to replicate some recent experiments on the comprehension of plurality in Turkish, which found that plurals under negation in Turkish behave more like plurals in English.

11 Addressing Daoxin Li's question.

6. References

Ackerman, F., Butt, M., & King, T. H. (1998). Constructions and mixed categories: Determining the semantic interpretation of person/number marking. In *Proceedings of the LFG98 Conference*.

Austin, J. (2013). Markedness, Input Frequency, and the Acquisition of Inflection: Evidence from Basque/Spanish Bilingual Children. *International Journal of Bilingualism*, 17(3), 259-283.

Bale, A., & Khanjian Hrayr (2008) Armenian Classifiers and Number Marking, *Proceedings of SALT* 18: 73-89.

Bale, A., Gagnon, M. & Khanjian, H. (2010). On the Relationship Between Morphological and Semantic Markedness: The Case of Plural Morphology, *Morphology* (2011) 21:197–221

Bale, A., Gagnon, M. & Khanjian, H. (2011). Cross-linguistic Representations of Numerals and Number Marking, *Proceedings of SALT* 20: 582–598.

Bale, A. (2013). *Number, Competition and Syntactic Complexity*, based off NELS 2011 handout, 14 pp.

Bale, A., & Khanjian H. (2014). Syntactic Complexity and Competition: The Singular-Plural Distinction in Western Armenian, Linguistic Inquiry, 45(1):1–26.

Bel, A., & Rosado, E. (2009). Person and Number Asymmetries in Child Catalan and Spanish. *Hispanic Child Languages*, 195-214.

Buda, J. (1992). Language shift in Australia and Canada. *Otsuma Women's University Annual Report: Humanities and Social Sciences*, 24.

Bylinina, Lisa & Nouwen, Rick (2018). On "Zero" and Semantic Plurality. *Glossa: A Journal of General Linguistics* 3(1): 98. 1–23.

Carey, S. & Barner, D. (2019). Ontogenetic Origins of Human Integer Representations. *Trends in Cognitive Sciences*, 23, 823-835.

Chahinian, T., & Bakalian, A. (2016). Language in Armenian American communities: Western Armenian and efforts for preservation. *International Journal of the Sociology of Language*, 2016(237), 37-57.

Corbett, Greville G. (2000). *Number*. Oxford University Press, 358 pp.

Deal, Amy Rose (2013). Apportionment and the Mass-Count Distinction in Nez Perce, unpublished paper, UC Santa Cruz, 13 pp.

Donabedian-Demopoulos, A. (2018). Middle East and beyond-Western Armenian at the crossroads: A sociolinguistic and typological sketch. In Bulut, Christiane, *Linguistic minorities in Turkey and Turkic-speaking minorities of the periphery*, 111/2018, Harrazowitz Verlag, 111, Harrazowitz Verlag, *Turcologica*, 89-148.

Filian, C. (2018). *The impact of residential location on language practices among Lebanese Armenian adolescents: an investigation into the Western Armenian language in Lebanon* (Doctoral dissertation, King's College London)

Francez I. & Koontz-Garboden, A. (2017). *Semantics and Morphosyntactic Variation, Oxford Studies in Theoretical Linguistics*, Oxford University Press, 192 pp.

Godson, L. I. (2004a). Vowel production in the speech of Western Armenian heritage speakers. *Heritage Language Journa*l, 2(1), 44-69.

Godson, L. I. (2004b). *Phonetics of language attrition: Vowel production and articulatory setting in the speech of Western Armenian heritage speakers*. (Doctoral dissertation, University of California at San Diego).

Hanson, R., Harley, H., & Ritter, E. (2000). Underspecification and Universal Defaults for Person and Number Features. In *Actes de l'ACL/CLA conference proceedings*. *Ottawa: Cahiers Linguistiques d'Ottawa*, 111-122.

Harley, H., & Ritter, E. (2002). Structuring the bundle: A universal morphosyntactic feature geometry. In *Pronouns: Features and representation*, 23-39.

Huang, Y., Spelke, E. S., & Snedeker, J. (2013). What exactly do number words mean? *Language Learning and Development*, 9(2), 105-129.

Karapetian, S. (2014). *"How do I teach my kids my broken Armenian?": A study of Eastern Armenian heritage language speakers in Los Angeles.* (Doctoral dissertation, UCLA).

Karimian, N. C. (2020). *Armenian-American parents' role in Armenian heritage language development.* (Doctoral dissertation, University of Southern California).

Kelly, N. E., & Keshishian, L. (2019). The Voicing Contrast in Stops and Affricates in the Western Armenian of Lebanon. In *Interspeech*, 1721-1725.

Landman, F. (2000). *Events and Plurality: The Jerusalem Lectures*, vol. 76. Boston, MA: Kluwer Academic Pub. doi, 10, 978-94.

Legendre, G. et al. (2010). Experimental and Empirical Evidence for the Status and Acquisition of Subject Clitics and Agreement Marking in Adult and Child Spoken French. *Movement and Clitics*, 333-360.

Lonsdale, D., & Danielyan, I. (2004). A two-level implementation for Western Armenian morphology. *Annual of Armenian Linguistics*, 24, 35-51.

Martí, L. (2017). Zero N: number features and⊥. Unpublished manuscript, Queen Mary University of London.

Martí, L. (2020). Numerals and the Theory of Number. *Semantics and Pragmatics*, 13(3), 1-54.

Matsumoto, Y. (1985). Acquisition of some Japanese numeral classifiers: the search for convention. *Papers and Reports on Child Language Development*, 24, 79-86.

Montrul, S. (2016). *The acquisition of heritage languages.* Cambridge University Press.

Paperno, D. (2012). *Semantics and Syntax of Non-Standard Coordination.* (Doctoral dissertation, UCLA).

Pereltsvaig, A. (2013). On Number and Numberlessness in Languages without Articles, *Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society (2013)*, pp. 300-314.

Potts, C. (2008). Review article: Hagit Borer's *Structuring Sense, Volumes I and II. Language* 84(2):343-369.

Ritter, E. (1991). Two functional categories in noun phrases: Evidence from Modern Hebrew. In *Perspectives on phrase structure: Heads and licensing*, 37-62. Brill.

Scontras, G. C. (2013) A Unified Semantics for Number Marking, Numerals, and Nominal Structure, *Proceedings of Sinn und Bedeutung*. Vol. 17, 545-562.

Scontras, G. C. (2014). *The semantics of measurement* (Doctoral dissertation, Harvard University).

Sigler, M. (1997). *Specificity and agreement in standard Western Armenian*. (Doctoral dissertation, Massachusetts Institute of Technology).

Sudo, Y. (2015). PLIN3005/PLING229 Advanced Semantic Theory B, lecture notes, *University College London*, 163 pp.

de Swart, H., & Farkas, D. (2010). The semantics and pragmatics of plurals. *Semantics and Pragmatics*, 3, 6-1.

Wolf, M. (2013). Candidate Chains, Unfaithful Spell-Out, and Outwards-Looking Phonologically-Conditioned Allomorphy, *Morphology*, 23:145–178.

Yamamoto, K., & Keil, F. (2000). The acquisition of Japanese numeral classifiers: Linkage between grammatical forms and conceptual categories. *Journal of East Asian Linguistics*, 9(4), 379-409.