The emergence of natural language quantification

Annemarie Kocab[1,2], Kathryn Davidson[2], & Jesse Snedeker[1]

[1]Department of Psychology, Harvard University

[2]Department of Linguistics, Harvard University

Annemarie Kocab (corresponding)

Department of Psychology

Harvard University

33 Kirkland St.

Cambridge, MA 02138

kocab@fas.harvard.edu

Abstract

Classical quantifiers (like “all”, “some”, and “none”) express relationships between two sets, allowing us to make generalizations (like “no elephants fly”). Devices like these appear to be universal in human languages. Is the ubiquity of quantification due to a universal property of the human mind or is it attributable to more gradual convergence through cultural evolution? We investigated whether classical quantifiers are present in a new language emerging in isolation from other languages, Nicaraguan Sign Language (NSL). An observational study of historical data collected in the 1990s found evidence of potential quantifier forms. To confirm the quantificational meaning of these signs, we designed a production study that elicited, from three age cohorts of NSL signers (N=17), three types of quantifiers: universal (all), existential (some), and negative (none). We find evidence for these classical quantifiers in the very first generation of signers, suggesting they may reflect a universal property of human cognition or a very rapid construction process.

Keywords: quantifiers, language emergence, language evolution, sign language, semantics

The emergence of natural language quantification

1. Introduction

All human languages have ways of expressing information about quantities. One of the most common is with quantifiers like “some” and “all” that relate two sets (“all babies sleep” conveys the relationship that babies are a subset of sleeping things). We call these expressions *classical quantifiers*. Noun phrases with quantifiers appear in the same syntactic positions as other noun phrases but they have a very different function. Referential noun phrases, like “those elephants” or “Joe Biden” allow us to make assertions about particular entities in the world. Classical quantifiers, in contrast, readily allow us to describe generalizations and hypotheses about the world, such as “Most bananas are clones,” “No elephants lay eggs”, and so on.

Classical quantifiers have played a central role in the history of the study of natural language, pushing semanticists to adopt more powerful logical systems (Barwise & Cooper 1981; Frege, 1879; Montague, 1973). As semantic operators, quantifiers like *some* and *all* relate two sets (e.g., the set of elephants and the set of egg-layers) but as syntactic determiners they take only one noun as their argument, thus they force us to more complicated models of the relation between meaning and form. Quantifiers have also been central to our understanding of human cognitive development because they provide insight into the relationship between children’s cognitive capacities and linguistic knowledge (Brooks & Braine, 1996; Inhelder & Piaget, 1964; Piaget, 1968).

All human languages appear to have devices like classical quantifiers, although there is cross-linguistic variation in how they are expressed (Bach, Jelinek, Kratzer, & Partee, 1995; Davis, 2010; Keenan & Paperno, 2017; Partee, 1995). As far as we know, nothing similar to classical quantifiers has been reported in any animal communication system. This raises the question of their origins. Are quantifiers present across languages because they encode concepts that are readily available to all humans regardless of the language that they use and thus can emerge in a single generation? Or do they arise over many centuries through convergent cultural evolution, perhaps due to universal communicative pressures? Or do they emerge on an intermediate time scale, in two or three generations, as learners reanalyze and refine the language they are exposed to? In this paper, we explore this question by examining whether quantifier expressions are present in a new emerging language, Nicaraguan Sign Language (NSL).

NSL is a newly emergent language, only around forty-five years old. Prior to the 1970s, there was no commonly used sign language in Nicaragua. In Nicaragua, as elsewhere, deaf children with no access to language models came up with gestural systems for communicating with their families, called homesign. Education reforms led to the opening of a school for special education in 1977, which allowed deaf children and adolescents to socialize in greater numbers than ever before (Kegl & Iwata, 1989; Polich, 2005). This initial group of fifty signers began developing a language which they passed on to waves of new children who entered the community each year. These children, in turn, added to the language's complexity (Senghas, 1995; Senghas & Coppola, 2001). By comparing the language of that initial group of signers (the first cohort) to that of those who entered in the language's second decade (the second cohort) and those who entered in the third decade (the third cohort) we can measure how the language has changed and grown over time.

The present study explores whether classical quantifiers are used by each of these cohorts. To date, there has been no work on this question in NSL or any other emerging language and we had no strong a priori hypotheses about when quantificational language would emerge in a new language. Thus, we considered three competing hypotheses based on arguments that have been proposed about the origins of language and relevant empirical findings. The first possibility is that classical quantification is a basic property of human cognition available to all learners prior to language acquisition, in which case we might expect quantificational language to be present even in a language's first stages. The second possibility is that classical quantifiers emerge over several generations as learners reanalyze their input. In this case, we might expect classical quantifiers to be present in the language of the third cohort. The third possibility is that classical quantification is historically constructed as languages mature, emerging only after significant time, and thus is absent in a young developing language. Note that the strongest test of the first possibility would be to investigate whether there is quantificational language in homesigners, individuals who have never acquired an external language, a point to which we return in the General Discussion. For practical reasons in our first investigation of quantificational language, we focused on the first three cohorts of NSL signers.

There are two reasons why we might expect to observe classical quantifiers in NSL. First, quantificational expressions (often as standalone words, or in some languages as parts of other words) are argued to be universal (Bach et al., 1995; Barwise & Cooper 1981; Partee, 1995 cf. Davis, 2010; Matthewson, 2014). One reason a linguistic property could be universal (or nearly universal) is because it reflects a stable property of the human mind (e.g., Chomsky, 1988; Christiansen & Chater, 2008; Pinker & Bloom, 1990).

Second, we might expect to find classical quantifiers in NSL because exact numbers, which have many of the same properties, appeared early. First-cohort signers of NSL have a set of systematic signs for expressing cardinal numbers (Flaherty & Senghas, 2011). Number words occur in many of the same syntactic contexts as classical quantifiers and also express quantity information. In many semantic theories, number words, like classical quantifiers, are analyzed as expressing the relation between two sets, such that “Two elephants wear hats” and “All the elephants wear hats” both describe a relation between the sets of elephants and hat-wearers (Barwise & Cooper, 1981; 1984).

There are, however, several reasons why we might not expect classical quantifiers to be present in NSL. First, as noted above, classical quantifiers have provided a central motivation for theories that posit a more complex interface between meaning and form because they express the relation between two sets but take only one noun as their syntactic argument. Developing a linguistic system capable of making this mapping may require cultural processes that unfold over many generations or repeated refinement via iterated learning. At first glance, this hypothesis might seem to be incompatible with the prior work on numbers in NSL. Like classical quantifiers, number words can also be used to relate two sets (“Two elephants flew kites”). Because the first cohort of NSL users have number signs, one might assume that they use these number words to express the relation between two sets in much the same way as classical quantifiers. The existing data, however, do not justify this assumption. Number words, unlike classical quantifiers, can and often are used to specify the cardinality of a single set. For example, the numbers in sentences like “There are two elephants” and “The two elephants walked into a bar” are typically analyzed as one place predicates with a semantic structure that matches their syntactic structure (see Ionin & Matushansky, 2006 for example). Prior research on number words in NSL was not designed to determine whether the number words were being used as one or two place predicates. For example, in Flaherty and Senghas (2011), the elicitation contexts involved a single set of objects (e.g., seven tokens) and a simple prompt (“How many?”) making it unclear whether the number signs that were produced (“seven”) should be construed as a one place predicates expressing the numerosity of that set (“There are seven tokens") or as a two place predicates expressing the intersection between that set and another predicate that the participant happened to have in mind ("Seven tokens are on the table").

Second, numbers could emerge earlier than classical quantifiers because they express properties that are more perceptually accessible or easier to learn. The cardinality of a small set can be quickly determined by looking at the reference set alone (unlike "allness" or "someness" which requires knowing about the complement set as well). Even very young infants have the ability to extract the cardinality of a set (e.g., twoness) and generalize it across sets of different kinds (e.g., "two horses" and "two balls") (e.g., Huntley-Fenner, Carey, & Solimando, 2002; Wynn, 1996). In contrast, the meanings of classical quantifiers are necessarily relational and consequently these terms can be used to describe groups of wildly different cardinalities (e.g., "some birds fly" and "some of my children ate eggs"). Thus, cardinality may be perceptually available in a way that classic quantificational concepts are not. This would be consistent with studies showing that numbers are processed more efficiently and interpreted differently than classical quantifiers in both young children and adults (e.g., Huang & Snedeker, 2009a/b; Huang, Spelke, & Snedeker, 2013).

Next, exact numbers may have emerged quickly in NSL because the deaf children who created it were brought together in a school. Formal education places a heavy emphasis on counting and exact numerosity, and thus the children may have had an urgent need to find some way to communicate about quantity. In contrast, classical quantifiers would seem to be less frequent targets of explicit elementary school instruction.

Finally, a system for representing exact numbers may emerge quickly in sign languages because it can capitalize on the iconicity of the one-to-one correspondence between an integer and the value to which it refers. Katseff and Senghas (2004) and Flaherty and Senghas (2011) report that that the earliest number signs in NSL (before a conventionalized set emerged) were highly iconic, with a one-to-one mapping between the number of extended fingers and the number of entities to which the sign referred. The possibility of one-to-one mapping is not available for classical quantifiers.

Note that if quantifiers are constructed slowly over historical time, then we would still need an account of the raw materials they are constructed from and how children recapitulate this historical achievement. Such a proposal would, in some sense, situate quantifiers as a product of the human mind. Nevertheless, there is a critical distinction between theories in which assembling these materials takes place over years, and ones in which it takes place over millennia.

In our investigation of quantificational language in NSL, we compare quantifiers of three different forces: universal, existential, and negative. A universal quantifier is used when the property that defines one set is satisfied by every member of another set (given any x in P, property Q is true). An existential quantifier is used when the property of one set holds for at least one member of the other set (for some x in P, property Q is true). Note that existential quantifiers can be used in a context that is satisfied by a universal quantifier (e.g., "some birds have wings" can describe a case in which all birds have wings), though such descriptions are typically dispreferred because they are under informative, a point to which we return to in the discussion. The reverse does not hold ("all birds have wings" cannot truthfully describe a case in which only some birds have wings). We use the term negative quantifier to refer to quantificational expressions that express the negative of an existentially quantified statement. In other words, this expression is used when it is not the case that there is a member in the domain set of which the property holds that defines the other set (it is not true that there is an x in P with the property Q, or no x in P has the property Q). These quantificational forces were targeted because crosslinguistic evidence shows that similar meanings are expressed across different languages (Horn 1989, 2012). We did not have any hypotheses about whether certain quantifier forces would emerge earlier or later in a new language.

Prior fieldwork has documented several candidate signs in NSL that might serve the function of classical quantifiers (see Fig. 3). We investigate the emergence of quantifiers in the language of three cohorts of NSL signers in two studies. The first is an observational study of historical data collected in the 1990s in which we looked for instances of these signs. The second is an experimental study designed to determine the meaning of possible quantificational forms in a controlled language elicitation task. Together, these studies allow us to look at the development of classical quantifiers in historical time (comparing data from the 1990s to data collected in 2015) and across apparent time (comparing data collected across the cohorts in 2015).

2. Study 1: Corpus data

We investigated quantifiers in a sample of historical video footage collected between 1993-1996 by Ann Senghas for her dissertation (Senghas, 1997; Senghas & Coppola, 2001). These data are some of the earliest corpus data of NSL available to the authors. They primarily consisted of signed narratives of two Mr. Koumal stories by single participants and spontaneous conversations between pairs of participants.

Note that these elicitations were collected for a different purpose and did not target quantifiers. The stories of Mr. Koumal are short animated non-verbal cartoons. Two cartoons were used for these elicitations: “Mr. Koumal Flies Like a Bird” (Studio Animovaného Filmu, 1969) and “Mr. Koumal Battles his Conscience” (Studio Animovaného Filmu, 1973). These stories were chosen by the previous researchers because work with users of other sign languages had shown that they elicit a high number of spatial devices (verbal agreement, verbs of motion and location, coreference). For instance, in “Mr. Koumal Flies Like a Bird,” Koumal uses chicken feathers to make wings for himself. After crashing into a mountainside, he makes Indian headdresses from the feathers to sell to children. In the spontaneous conversations, participants were recorded having natural, everyday conversations. They were not asked to discuss any particular topic. The content of these conversations included personal anecdotes, discussions of current family dynamics, reminiscing about memories of school. As with any corpus analysis, the challenge of looking for particular structures is that those structures may be absent because they are low frequency and generally hard to see, the context does not support use of the particular structure, or the structure is not in the participant’s repertoire.

Moreover, studying the production of quantifiers in a given language faces the challenge that languages choose to lexicalize quantificational meaning in different ways, for example as adverbials (always) versus determiners (every) versus bound morphemes, or even adjacent meanings like generics, etc., so there is no particular ratio of purely quantificational vs. non-quantificational expressions we expect in a given language (Bach, Jelinek, Kratzer, & Partee, 1995; Keenan & Paperno, 2017). However, crosslinguistic comparisons show stable evidence of the existence of similar meanings expressed: human languages can express existential (e.g. some), universal (e.g. every, all), and negative existential (e.g. none = not some) meanings in single forms, while negative universal forms are not lexicalized (e.g. *nall) (Horn 1989, 2012). Thus these quantifiers represent a good target for our investigations.

Across approximately 7 hours of historical footage coded for quantificational expressions, we observed six different lexical signs used by 21 NSL signers to convey quantificational information: ALL, MANY, SOME, LITTLE, NONE-1 and NONE-2. These glosses are our English translations of the NSL signs based on their use in context. Nineteen of the NSL signers were first-cohort signers and two were young second-cohort signers. There were a total of 34 instances of possible quantificational signs: one universal, 24 existential, and 9 negative. All of these signs were embedded in a larger utterance. At least one instance of each quantifier type was used by a first-cohort signer. We provide three examples of the sentences with quantifiers we observed with the English translation in parentheses.

1. IX ALL FEATHERS PUT-ON (He put on all the feathers).
2. I LEARN LITTLE (I learned a little bit [in school]).
3. IX MONEY POCKET NO (He had no money in his pocket).

We note four things from these observations. First, these corpus data suggest that universal, existential, and negative quantifiers were used by the first group of NSL signers when they were young adults suggesting that these signs emerged early in the creation of NSL. Second, universal quantifiers, specifically ALL, were less frequent than existential and negative quantifiers in our sample. Third, there is often ambiguity in the intended domain of quantification (e.g., what defines the total set of available feathers) and in whether a quantifier relates two sets (e.g., I learned a little of the knowledge available) or simply describes the rough cardinality of a single set (e.g., the size of my knowledge was small). Finally, like all exercises in rich interpretation (Bloom, 1970), our glosses of these forms were informed by our belief that the signers were trying to convey thoughts similar to those that we might express, using conceptual building blocks that would be familiar to us. It is difficult, perhaps impossible, to tease apart our knowledge from that of our informants. Thus an experimental study is necessary: one in which the context is controlled (to provide validation for our interpretation of the message) and where the domain is clearly defined. We next describe such an experimental study.

3. Study 2: Experimental elicitation of quantifiers

An observational study based on data collected in the 1990s found evidence of lexical forms that appear to function like classical quantifiers used by first-cohort signers. This suggests that quantificational language may have emerged early in the development of NSL. However, we are limited in the kinds of inferences we can make on the basis of observational studies. An experimental study allows us to constrain our interpretation. Based on the observational findings it is possible that quantifiers appear in a stable form very early in a language's development, and so we might expect to observe no differences across apparent time between the cohorts in their quantificational language. Alternatively, an experimental study may allow us to (a) uncover differences in how the cohorts use quantifiers of different types and/or (b) document additional developments in the language with the second and third cohort of signers. For instance, given the relative infrequency of universal quantifiers in the early narratives, we might expect to see some differences across cohorts in the use of these forms. Given that the narratives were not designed to elicit quantificational language, we may also see other quantifier signs in addition to the forms we observed in Study 1.

4. Methods

4.1. Participants

The participants in this study were 17 deaf Nicaraguan signers who were exposed to the emerging sign language by the age of 6 years. Five entered the signing community before 1983 and are referred to as first-cohort participants (mean age: 43, mean age of entry: 4.7). Five entered between 1986 and 1990, and are referred to as second-cohort participants (mean age: 32, mean age of entry: 4.2). Seven entered between 1993 and 1999, and are referred to as third-cohort participants (mean age: 24, mean age of entry: 3.5). These data were collected in 2015, approximately 20 years after the narratives in Study 1.

Due to the unique phenomenon of NSL, our sample size is constrained. Prior research on NSL shows large effect sizes for cohort differences (w =.51-.68; Kocab, Senghas, & Snedeker, 2016). Note that these effect sizes are drawn from investigation of a different domain (temporal language), and it is possible that effect sizes for cohort differences vary across linguistic domains. Due to the relatively limited work on NSL and the lack of prior work on classical quantifiers in NSL specifically, we are unable to calculate effect sizes for quantificational language. However, the findings from past work in other domains suggests that despite small sample sizes, we have sufficient power to detect marked cohort differences if they are present (see section 5.1 for power analysis).

4.2. Materials

Each participant saw 12 sets of pictures presented in a Keynote presentation on a Macintosh laptop. In a pretest, the pictures were presented to 10 English-speaking participants on Amazon Mechanical Turk, where they elicited classical quantifiers.

On each trial, participants first saw one picture (e.g., a scene with bears standing on the shore of a lake). On the next slide, the first picture was greyed out and two new pictures appeared below. The two pictures depicted contrasting scenes, where different proportions of the set were engaged in a new activity thematically appropriate for the scene (e.g., some of the bears are swimming vs. all of the bears are swimming). The top picture never showed any member of the set engaged in this new activity. One of the two bottom pictures, the target, was outlined in red. Participants were asked to describe the target (Fig. 1).

The contrasting scenes were designed to elicit descriptions containing quantifiers of the three forces of interest: universal, existential, and negative. Four minimally-different versions of the contrasting pictures were created for each base event, in which the proportional differences represented in the picture varied. The universal quantifier picture depicted the entire set engaged in the new activity (e.g., a description in English would be "all of the bears are swimming"). The negative quantifier picture depicted none of the individuals engaged in the new activity (e.g., "none of the bears are swimming"). There were two different kinds of existential quantifier pictures: one with less than half of the individuals in the set engaged in the new activity (e.g., "some of the bears are swimming") and one with more than half (approximately three-quarters) of the individuals in the set engaged in the new activity (e.g., "most of the bears are swimming"). These were analyzed together (as different examples of existential force) but also separately (to explore whether a distinction is made between smaller and larger proportions).

The key difference between the pictures was the size of the subset engaged in the new activity. This contrast served to highlight the relevance of the proportional differences. The use of contrasting pictures that depicted a new activity (rather than showing three pictures of different sized sets engaged in the same activity) was by design to elicit descriptions that relate two sets rather than descriptions that could be interpreted as describing a single set.

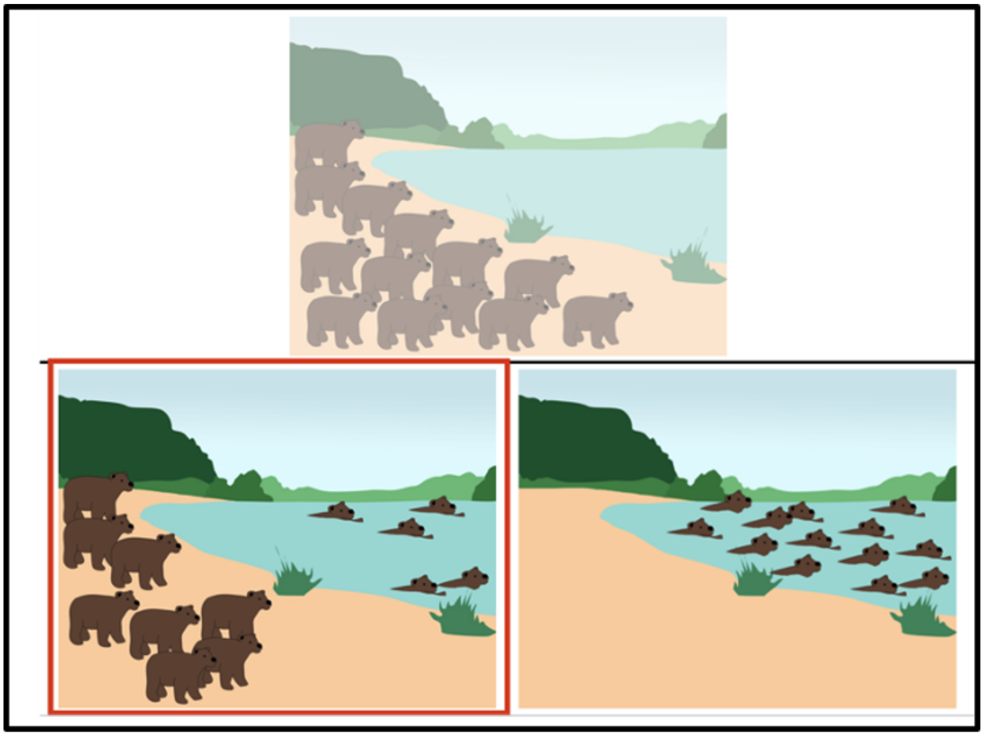

Figure 1. Sample stimulus item. The top picture depicts an initial scene of a group of bears standing by the shore of a lake. The bottom two pictures depict two contrasting scenes, designed to elicit descriptions with different quantifiers ("Some of the bears are swimming" for the target, left picture vs. "All of the bears are swimming" for the right picture). One of the two bottom pictures was outlined in red and designated the target picture.

One challenge for eliciting descriptions containing classical quantifiers using pictures is that any given scene can also be described with exact numbers (e.g., "five bears" rather than "some bears"). We took two steps to reduce the likelihood that participants would use numbers. First, each participant received four training trials at the beginning of the study, designed to elicit quantity descriptions containing mass nouns, such as "The girl drank all of the water" (Fig. 2). These trials were intended to nudge participants to construe the pictures as illustrating proportional relationships rather than exact numerosities. Second, to make counting more laborious and less convenient, all test pictures depicted large sets of animate entities (e.g., bears, boys, nurses), containing between 13 to 19 individuals.

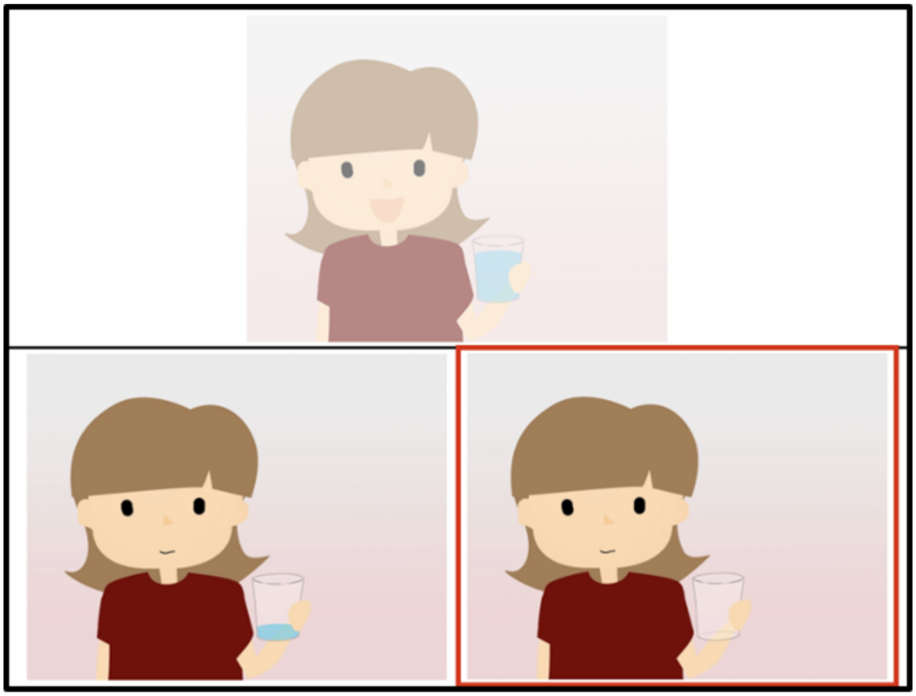

Figure 2. Sample training item. The top picture depicts an initial scene of a girl holding a full glass of water. The bottom two pictures depict two contrasting scenes, designed to elicit quantificational expressions containing mass nouns describing different quantities ("The girl drank most of the water" for the left picture vs. "The girl drank all of the water" for the right picture).

Two lists were created, and approximately half of the participants received each list. Each list consisted of the same 12 base events. The target quantificational force and alternative for each trial were randomly generated separately for each list with the constraint that no more than two sequential trials had the same quantifier picture type as the target. Thus, each of the four quantifier picture types described above was the target three times in each list. The side of the screen, left or right, that the target photo appeared on was counterbalanced across trials.

4.3. Procedure

Participants were told that they would see a before scene and then two different after scenes. They were asked to describe the after scene that was outlined in a red box, indicating the target picture. They were not prohibited from describing the other picture, and indeed many participants spontaneously provided descriptions of both after scenes. If the participant provided a description with an exact number, the experimenter asked if there was another way to describe that picture without counting or using numbers[1]. If the participant failed to provide a quantifier term in the initial description, the experimenter asked how the pictures were different from each other and whether the creatures were engaged in a particular action (e.g., the experimenter might ask if there are bears swimming). If the participant still did not provide a description with quantity information, the experimenter moved on to the next trial. Over 75% of the trials were described with a lexical quantifier (238 out of 312 trials).

4.4. Coding

The data were coded offline, by the first author who has been working with NSL signers for over a decade, using the ELAN video and audio annotation program (http://tla.mpi.nl/tools/tla-tools/elan/) developed by the Max Planck Institute for Psycholinguistics (Lausberg & Sloetjes, 2009; Wittenburg, Brugman, Russel, Klassmann, & Sloetjes, 2006). A second coder who was blind to cohort coded 35% of the descriptions (6 of 17 participants). Interrater reliability across variables of interest was 85%.

The data were coded for the presence of three devices: classical quantifiers, numbers, and a third category used by many participants which we will call verbal quantity expressions.

The first device, classical quantifiers, was the category that our stimuli were most straightforwardly intended to elicit. In this dataset, we observed eight lexical signs used by NSL signers to convey quantificational information: ALL, MANY, SOME, HALF, LITTLE, NONE-1, NONE-2, and NONE-3. Note again, that these glosses are our working hypotheses about the

[1] An English gloss of the instruction given in NSL is: Let's put aside numbers and counting. How are the top picture and the bottom picture different from each other? A rough sign-by-sign gloss of this is: NUMBER COUNTING PUT-ASIDE. $\text{IX}_{\text{top picture}}$ $\text{IX}_{\text{target picture}}$ DIFFERENT HOW?

approximate English translations of the NSL signs. Systematic use of these forms in the relevant contexts would support these analyses. Based on our field observations, we identified three different NSL signs for the concept NONE, each with a different form. All three forms were treated as negative quantifiers. We did not observe any systematic differences in the use of the different forms of NONE based on cohort or context.

Figure 3. Examples of classical quantifiers. Left image: MANY. Center left image: HALF. Center right image: LITTLE. Right image: one form of NONE.

The second device we coded was the use of exact cardinal numbers, since signers may still produce some numbers despite our attempt to limit them in the design.

The third device, verbal quantity expressions, has been documented in spoken languages (e.g., Gil, 1993) and sign languages (e.g., Kimmelman, 2017; Petronio, 1995). For example, in American Sign Language (ASL), verbs can take one morpheme to describe a set of one and a different morpheme to describe a set of many[2] (Petronio, 1995). For example, in the sentence STORE, MAN GO-TO, the sign for the verb GO-TO can be produced with a single extended finger to mean one man went to the store (unlike in English, plurality is not marked on the noun). A different handshape, representing a group of multiple individuals without specifying the exact number, can be used to mean that many men went to the store. An unmarked verb (produced

[2] In ASL, verbs can also contain different morphemes for sets of 1 to 4 (Petronio, 1995). As all of the sets in our stimuli contained more than five individuals (the number of digits on one hand), we did not expect to find (and indeed, we did not observe any) morphemes for quantificational values of 1-4 or 5.

without a number or plural morpheme) would be unspecified for number. The verbs in this dataset were coded as uninflected or containing a morpheme marking plurality (Fig. 4).

Note that this coding of verbs as uninflected or as containing a plural morpheme is based on our best interpretation of the forms. To our knowledge, there has been no systematic investigation of number classifiers in NSL. If the verbs appeared in their citation form, they were coded as uninflected. The verbs with a plural morpheme in this dataset (most frequently GO, FLY, and CLIMB) were produced with one hand, generally with all five fingers extended (see Fig. 4). While it is possible that this handshape is intended to represent the number five, that does not consistently map onto the number of the subset in the stimuli which varied.

If the participant provided multiple descriptions of the same picture, with and without use of a device, the description that contained the device was included in the analysis and the device was coded as present for the trial. If the participant provided a description of the target and the non-target picture containing one or more devices, both descriptions were included in the analysis. If a participant produced a description containing multiple devices, each device was coded as present for that trial.

Figure 4. The top series of panels depicts an uninflected form of the verb GO, not marked for number. The bottom series of panels depicts an inflected form of the verb GO, with plural marking. The left panel shows the starting point of the verb and the right panel the ending point. In the uninflected form, the sign begins with the signer's fingers grouped together. In the inflected form, the sign begins with the signer's fingers spread apart.

5. Results

5.1. Data analysis

Trials were excluded if the participant was not clearly visible (e.g., the camera was occluded by a passerby) or if it was unclear which picture the participant was describing. Two trials were excluded on the basis of these criteria.

Using the lme4 package in R (Bates, Maechler, Bolker, & Walker, 2014), analyses were conducted using generalized logistic mixed-effects models (Baayen, Davidson, & Bates, 2008).

For all models, item and subject were entered as random effects with random intercepts. We did not include a full random structure as we did not want to overfit the model. The presence of each dependent variable of interest (e.g., use of classical quantifiers, the use of multiple devices) was entered as a 1, and its absence as a 0[3]. Cohort and condition (quantifier force type) were each coded using two dummy variables. For analyses looking at cohort, the first cohort served as the baseline (the intercept). For analyses looking at condition, the Existential context served as the baseline. If a model failed to converge, we used the Nelder Mead and optimx optimization packages to improve model fit (Bihorel & Baudin, 2018; Nash, Varadhan, & Grothendieck, 2018).

Using the SIMR package and powerSim function, we have determined that with large effect sizes, we are sufficiently powered to detect effects of both cohort and condition via model comparisons (see above). For cohort effects, our observed power is in the range of 88.00% (95% Confidence Interval = 79.98, 93.64%). For condition, our observed power is in the range of 100.00% (95% Confidence Interval = 96.38, 100.00%).

We conducted the following analyses, looking at how often different quantificational devices are used (section 5.2), which classical quantifiers are produced in which contexts (section 5.3), and whether signers produce different classical quantifiers in the existential contexts based on the proportion of the set (section 5.4).

5.2. Use of different quantificational devices

First, we compared the use of the three devices for expressing quantification across the cohorts (Fig. 5). For each device, a model with only random effects and a model with the random

[3] The use of multiple devices was analyzed as a separate variable. If one or more devices were used, the variable was entered as a 1 for that trial. If one or no devices were used, the variable was entered as a 0.

effects and cohort as an independent variable were compared using the ANOVA function. Across all conditions, signers from all three cohorts produced classical quantifiers on the majority of trials (Cohort 1: 81%, Cohort 2: 82%, Cohort 3: 70%). We found no effect of cohort for the use of classical quantifiers [$\chi^2(2) = 2.26$, p=.324], the use of numbers [$\chi^2(2) = 1.92$, p=.384], or the use of verbal quantity expressions [$\chi^2(2) = 0.59$, p=.744]. We also found no difference across cohorts in the use of multiple devices [$\chi^2(2) = 0.53$, p=.766].

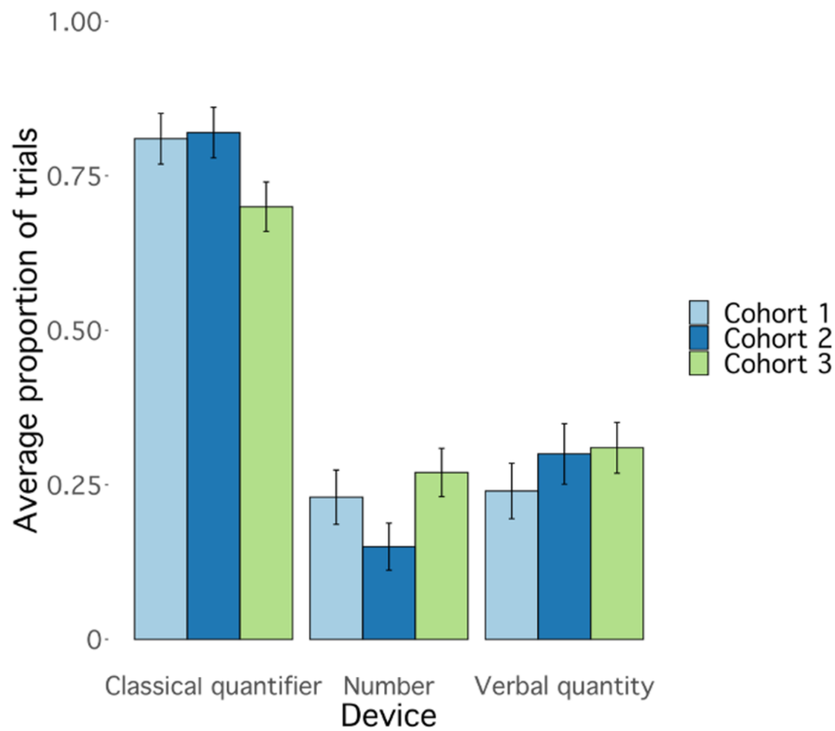

Figure 5. Proportion of trials in which a given linguistic device was used. Note that participants may have used multiple linguistic devices on a single trial.

5.3. Use of classical quantifiers across contexts

Our central questions concerned the distribution and quantificational force of the classical quantifiers in the three conditions (existential, negative, and universal) and whether this changed across the cohorts. For this analysis, because we had no strong hypothesis about differences in the use of existential forms, we collapsed the existential trials with more than half engaged in the action with and those with less than half. For each context, we constructed a series of four models: the random effects (base) model, a model with condition as a predictor variable, a model with condition and cohort as predictor variables, and a model with the interaction of cohort and condition. Using the ANOVA function, we compared each model with the next more complex model.[4]

First, we compared how often any classical quantifier was used in the three conditions, universal, existential, and negative context trials (Fig. 6). We found an effect of condition (*p*=.001, see Table 1 for model comparisons and Table 2 for effect sizes comparison), but no effect of adding the cohort fixed effect to the condition model (*p*=.316) and no effect of adding the interaction (*p*=.532). Specifically, signers produced more classical quantifiers in the universal and negative contexts than in the existential contexts (see Fig. 6) and this pattern was present in all three cohorts. Critically, in all of the cohorts, more than half of the responses in each condition involved a classical quantifier.

Dependent variable	Model 1	Model 2	(df) χ^2	p-value
Use of Classical Quantifiers	Random	Condition	(2) 13.99	.001
	Condition	Condition+Cohort	(2) 2.31	.316
	Condition+Cohort	Condition*Cohort	(4) 3.16	.532

Table 1. Model comparisons for the mixed-effects logistic regression predicting the use of classical quantifiers across condition and cohort.

[4] In many cases, the odds ratios obtained from our models are quite large, reflecting the small number of participants in our study and the large size of the effects (i.e., the strong categorical response patterns, see Figures 7, 9, 10). Given that we were unable to determine some effect sizes using mixed models, we conducted simpler subject based analyses, collapsing across items. We report these results in the Supplementary Materials for reference.

Dependent variable	Model	Predictor	Odds ratio	95% CI
Use of Classical Quantifiers	Condition ($p < .001$)	Intercept		
		Negative quantifier	2.299	1.081, 4.887
		Universal quantifier	3.923	1.730, 8.896
	Condition + Cohort (p=.316)			
	Condition * Cohort (p=.532)			

Table 2. Effect size estimates for the best fit model predicting the use of classical quantifiers across condition and cohort.

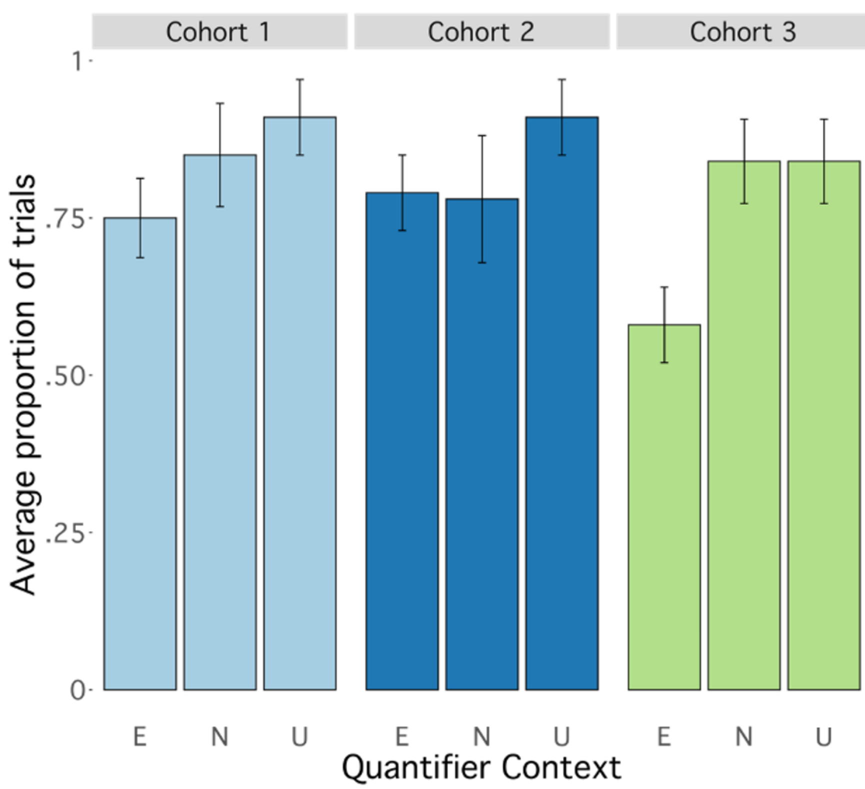

Figure 6. Proportion of trials in which a classical quantifier was used across existential (E), negative (N), and universal (U) Contexts.

5.3.1. Description of classical quantifier uses in existential, negative, and universal contexts

Next, we analyzed the classical quantifiers produced in each of the three contexts. First, we looked at the classical quantifiers produced in the existential contexts (Fig. 7). On the existential trials, when a classical quantifier was produced, signers from all three cohorts consistently produced forms that we had identified as likely existential quantifiers (Cohort 1: 100%, Cohort 2: 100%, Cohort 3: 100%).

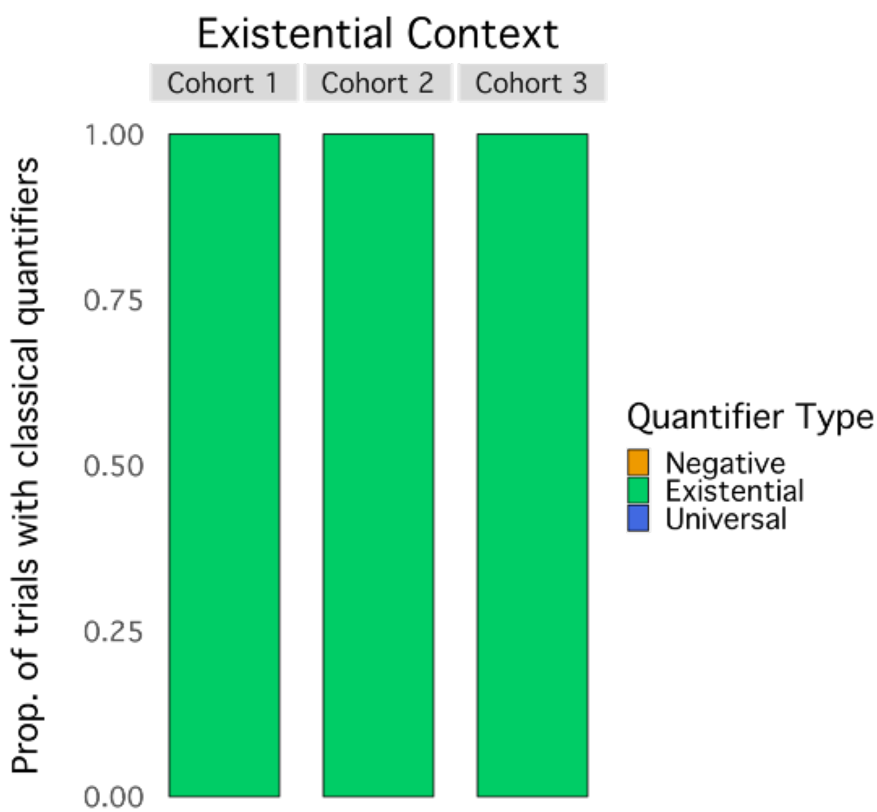

Figure 7. Type of quantifier produced on existential context trials when a classical quantifier was produced. Signers from all cohorts produced only existential quantifiers (e.g., MANY, SOME) in existential trials.

Then, we looked at the negative contexts (Fig. 8). On the negative trials, when a classical quantifier was produced, signers from all three cohorts predominantly produced negative quantifiers (Cohort 1: 71%, Cohort 2: 79%, Cohort 3: 85%). When a quantifier of a different type was produced in this context, it was paired with a different predicate (e.g., "All of the bears are standing" instead of "None of the bears are swimming").

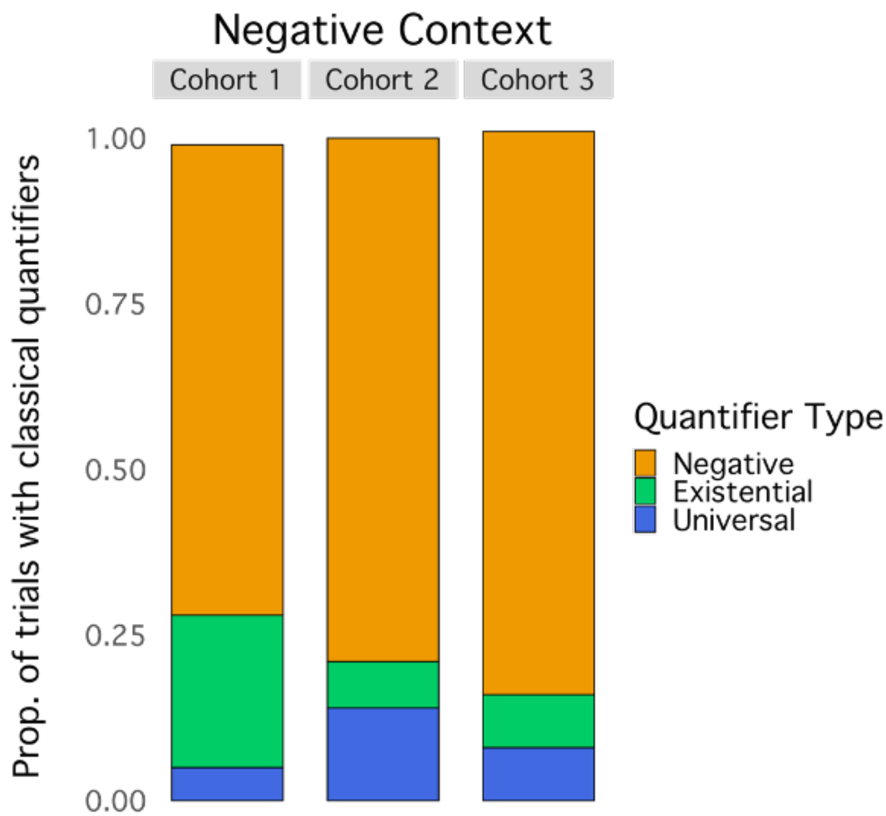

Figure 8. Type of quantifier produced on negative context trials when a classical quantifier was produced. Signers from all cohorts produced primarily negative quantifiers (e.g., NONE) in negative trials. Universal and existential quantifiers produced in this context were paired with a different predicate (e.g., "All of the bears are standing" instead of "None of the bears are swimming.")

Finally, we looked at the universal contexts (Fig. 9). On the universal trials, when a classical quantifier was used, second- and third-cohort signers produced mostly universal quantifiers (Cohort 2: 81%, Cohort 3: 69%) whereas first-cohort signers produced these quantifiers less frequently (Cohort 1: 24%) and often used existentials instead. Across all cohorts, of the 77 descriptions produced in a universal context, 68 occurred with a lexical quantifier. Of those 68 descriptions, 40 were described with a universal quantifier, 27 with an existential quantifier, and one with a negative quantifier. When an existential classical quantifier was produced in a universal context, it was nearly always the quantificational sign MANY (26 of 27 instances).

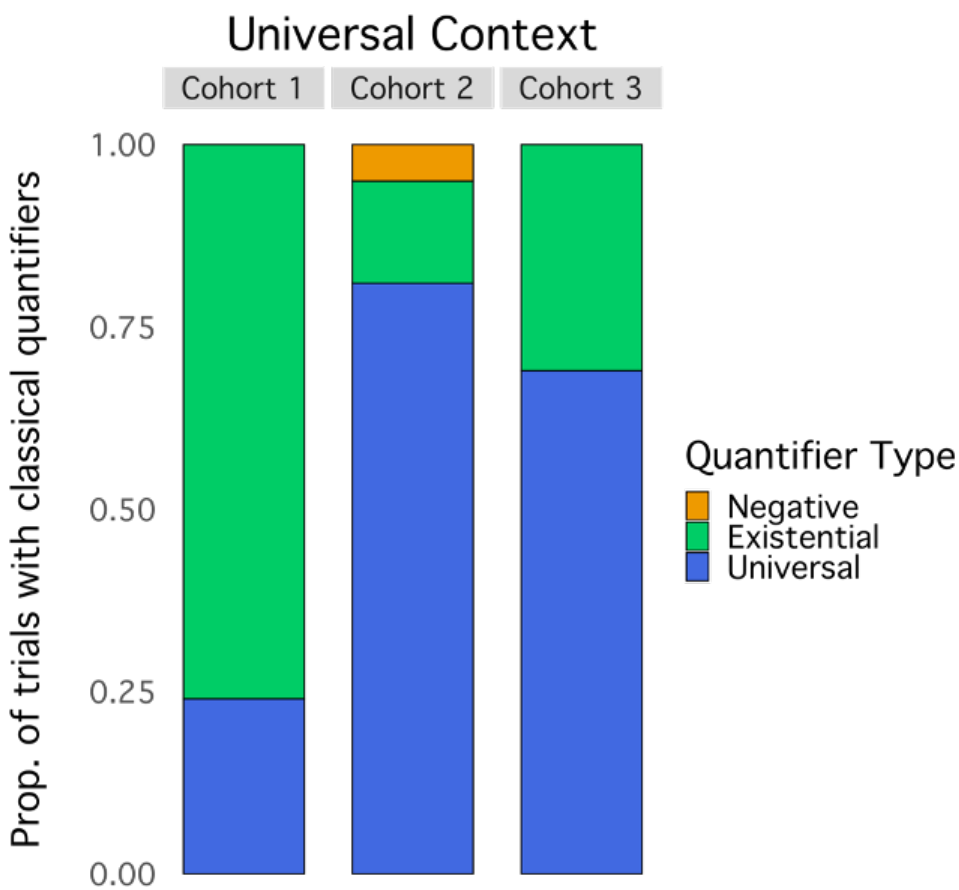

Figure 9. Type of quantifier produced on universal context trials when a classical quantifier was produced. Signers from the second and third cohorts produced primarily universal quantifiers (ALL) in universal trials. Signers from the first cohort produced primarily existential quantifiers (e.g., MANY). Negative quantifiers produced in this context were paired with a different predicate (e.g., "None of the bears are standing" instead of "All of the bears are swimming.")

We next conducted model comparisons of how the different classical quantifier forms are used across contexts.

5.3.2. Statistical analyses of existential classical quantifiers use across contexts

We constructed models to compare the use of existential quantifiers across conditions and cohorts. We found an effect of condition (*p*<.001, see Table 3 for model comparisons and Table 4 for effect sizes). Signers from all three cohorts consistently produced existential quantifiers in the existential contexts (see Fig. 7). There was also an effect of adding the cohort fixed effect to the condition model (*p*=.019) and an effect of including the interaction of condition and cohort (*p*=.009). These significant effects of cohort and interaction of condition and cohort are reflected in the first cohort's greater use of existential classical quantifiers in the Universal condition (see Fig. 10).

Dependent variable	Model 1	Model 2	(df) χ^2	p-value
Use of Existential Quantifiers	Random	Condition	(2) 66.98	<.001
	Condition	Condition+Cohort	(2) 7.94	.019
	Condition+Cohort	Condition*Cohort	(4) 13.51	.009

Table 3. Model comparisons for the mixed-effects logistic regression predicting the use of existential quantifiers across condition and cohort.

Dependent variable	Model	Predictor	Odds ratio	95% CI
Use of Existential Quantifiers	Condition	Intercept		
		Negative quantifier	0.035	0.013, 0.094
		Universal quantifier	0.195	0.096, 0.399
	Condition + Cohort	Intercept		

		Cohort 2	0.402	0.154, 1.05
		Cohort 3	0.246	0.100, 0.605
	Condition * Cohort	Intercept		
		Negative:Cohort 2	0.242	0.018, 3.197
		Universal:Cohort 2	0.038	0.006, 0.251
		Negative:Cohort 3	0.736	0.093, 5.793
		Universal:Cohort 3	0.274	0.060, 1.272

Table 4. Effect size estimates for the best fit model predicting the use of existential quantifiers across condition and cohort.

5.3.3. Statistical analyses of negative classical quantifiers use across contexts

We next conducted model comparisons examining the use of negative classical quantifiers. There was an effect of condition (p<.001, see Table 5 for model comparisons and Supplementary Materials for effect sizes) and no effect of including the cohort fixed effect to the condition model (p=.752). On the negative trials, participants in all three cohorts consistently produced negative quantifiers and there was no significant difference in the use of negative quantifiers by cohort.

Dependent variable	Model 1	Model 2	(df) χ^2	p-value
Use of	Random	Condition	(2) 149.84	<.001

Negative Quantifiers	Condition	Condition+Cohort	(2) .571	.752
	Condition+Cohort	Condition*Cohort	(4) 2.43	.657

Table 5. Model comparisons for the mixed-effects logistic regression predicting the use of negative quantifiers across condition and cohort. NB: Due to our small sample sizes and largely categorical response patterns, not all effect sizes could be calculated for the mixed effects logistic regression models. In order to calculate effect sizes, we ran alternative statistics, which yielded a similar pattern of results and enabled calculation of effect sizes. See Supplementary Materials for details.

5.3.4. Statistical analyses of universal classical quantifiers use across contexts

Finally, we conducted model comparisons looking at the use of universal classical quantifiers. We found an effect of condition (*p*<.001, see Table 6 for model comparisons and Supplementary Materials for effect sizes), an effect of adding the cohort fixed effect to the condition model (*p*=.039), but no effect of adding the interaction (*p*=.784). Participants in the second and third cohort used universal quantifiers frequently in the universal contexts and occasionally in the negative contexts, while the first cohort used them far less often (see Fig. 9).

Dependent variable	Model 1	Model 2	(df) χ^2	p-value
Use of Universal Quantifiers	Random	Condition	(2) 108.14	<.001
	Condition	Condition+Cohort	(2) 6.51	.039
	Condition+Cohort	Condition*Cohort	(4) 1.74	.784

Table 6. Model comparisons for the mixed-effects logistic regression predicting the use of universal quantifiers across condition and cohort. See Supplementary Materials for effect sizes.

5.4. Use of different classical quantifiers in existential contexts for different proportions

In the previous analyses looking at the use of classical quantifiers across the different contexts, we collapsed across the two existential context types (those with less than half of the individuals engaged in the new activity and those with more than half of the individuals engaged in the new activity). We next analyzed the two contexts separately to explore whether there are differences in the classical quantifiers produced for smaller and larger proportions.

We observed four existential quantifier signs produced in the two existential contexts: MANY, HALF, SOME, and LITTLE/FEW (Figs. 10-11). Across the three cohorts, MANY and LITTLE/FEW were the most frequently used quantifier signs. For the existential contexts that depicted more than half of the individuals engaged in the new activity ("Most"), on trials when a classical quantifier was produced, signers from all three cohorts primarily produced the quantificational sign MANY (Cohort 1: 100%; Cohort 2: 81%; Cohort 3: 60%, Fig. 10) and seldom produced LITTLE/FEW (Cohort 1: 0%; Cohort 2: 14%; Cohort 3: 5%). In contrast, for the existential contexts that depicted less than half of the individuals engaged in the new activity ("Some"), on trials when a classical quantifier was produced, signers produced fewer instances of MANY (Cohort 1: 50%; Cohort 2: 31%; Cohort 3: 15%, Fig. 11), and more instances of LITTLE/FEW (Cohort 1: 50%. Cohort 2: 44%; Cohort 3: 30%).

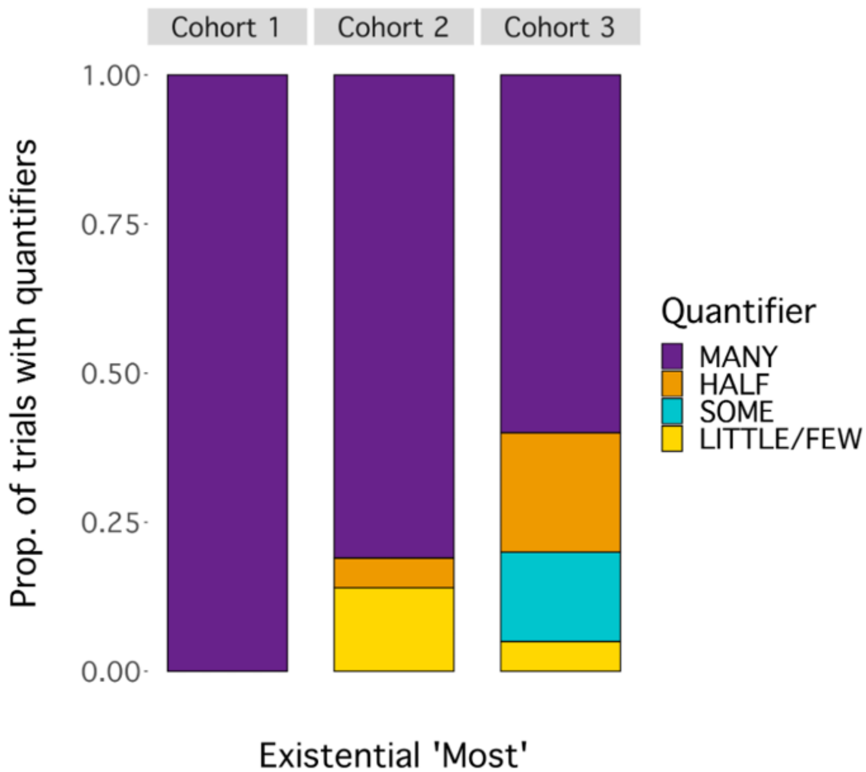

Figure 10. Existential quantifier signs produced in the “Most” existential context on trials when a classical quantifier was produced.

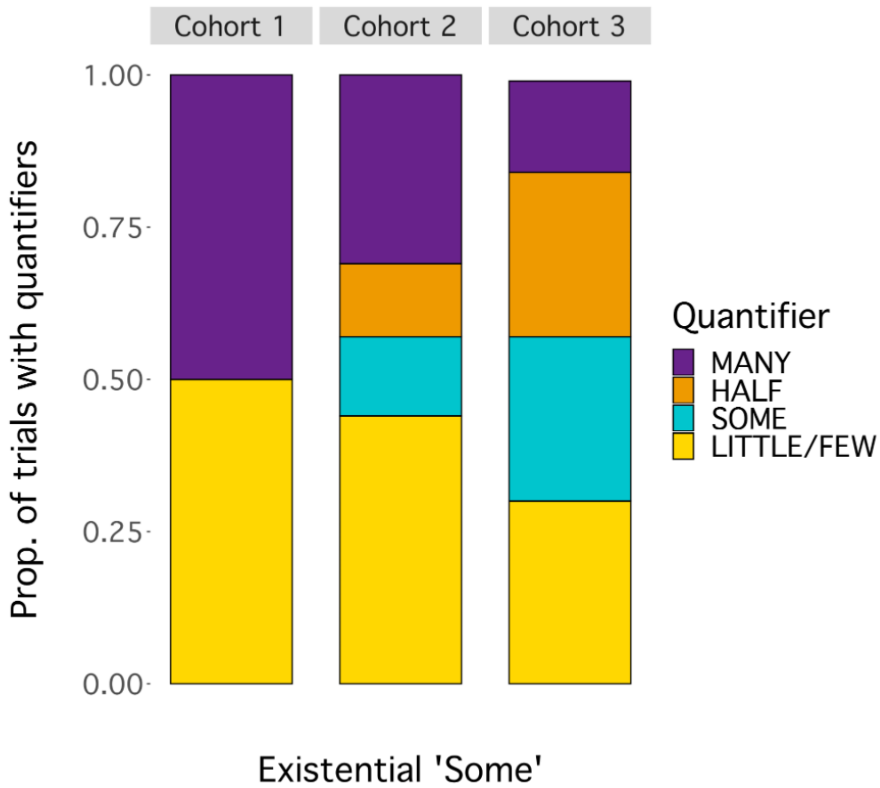

Figure 11. Existential quantifier signs produced in the "Some" existential context on trials when a classical quantifier was produced.

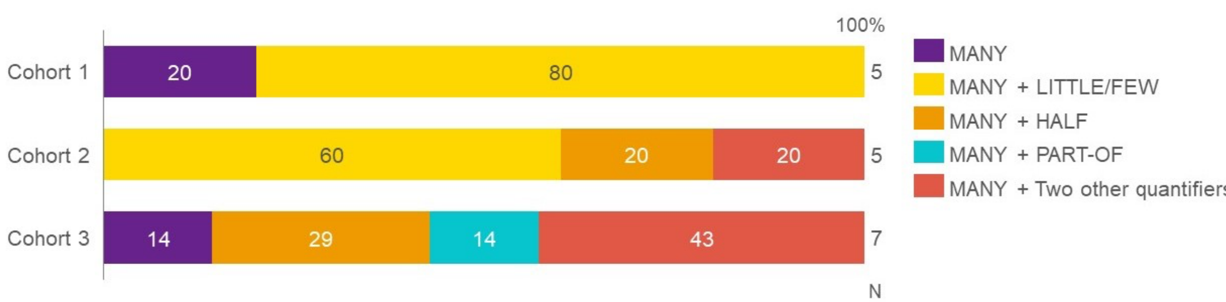

Figure 12. Use of existential quantifier signs in existential contexts by cohort. Colors represent different combinations of existential quantifiers used by signers.

Given the relatively infrequent uses of HALF and SOME (see Figs. 10-11), we focused our model comparisons on the classical quantifiers MANY and LITTLE/FEW in the existential contexts. The first model comparison looked at how often MANY was used and showed an effect of context (*p*<.001, see Table 7 for model comparisons). Signers from all three cohorts frequently produced MANY in the "Most" context and less so in the "Some" context. There was an effect of adding the cohort fixed effect to the context model (*p*=.047), but no effect of adding the interaction (*p*=.293). Older signers produced MANY in the "Most" contexts more frequently than the younger signers who, in addition to MANY, also used other classical quantifiers like HALF and SOME. In addition, compared to older signers, younger signers produced MANY less frequently in the "Some" contexts, instead producing more descriptions with HALF and SOME.

Dependent variable	Model 1	Model 2	(df) χ^2	p-value
Use of MANY	Random	Context	(1) 27.97	<.001
	Context	Context+Cohort	(2) 6.13	.047
	Context+Cohort	Context*Cohort	(2) 2.45	.293

Table 7. Model comparisons for the mixed-effects logistic regression predicting the use of the existential quantifier MANY across the two existential contexts and cohort.

The second model comparison looked at how often LITTLE/FEW was used in the two existential contexts and revealed an effect of context (*p*<.001, see Table 8 for model comparisons). Signers from all three cohorts frequently produced LITTLE/FEW in the "Some" context and less frequently in the "Most" context. There was no effect of adding the cohort fixed effect to the context model. Across all cohorts, LITTLE/FEW was used by signers more frequently in the "Some" contexts and less frequently in the "More" contexts.

Dependent variable	Model 1	Model 2	(df) χ^2	p-value

Use of SOME	Random	Context	(1) 20.62	<.001
	Context	Context+Cohort	(2) .52	.770

Table 8. Model comparisons for the mixed-effects logistic regression predicting the use of the existential quantifier SOME across the two existential contexts and cohort.

Together, these results suggest the beginnings of a differentiation in the classical quantifiers produced for smaller and larger proportions.

5.5. Discussion

The results of our experimental elicitation aligned with the observations of the historical data from the 1990s. We found quantificational language in the early stages of NSL, including dedicated lexical forms that functioned like classical quantifiers. Most of the signers (14 out of 17) produced *at least* one instance each of universal, existential, and negative lexical quantifier signs. (Two signers, one from the first cohort and one from the third cohort, produced only existential and negative quantifiers and one first-cohort signer produced only existential quantifiers.)

These signs were systematically produced in different contexts, validating our interpretation of the forms. Specifically, in existential contexts, all signers in all cohorts produced existential forms. In negative contexts, signers in all cohorts showed a strong preference for negative forms: all the apparent errors (existential or universal forms) were used with predicates that clearly indicated that the participants had a different construal of the picture in mind (e.g., all of the bears are standing vs. none of the bears are swimming). Only in the case of universal contexts did we see systematic differences across the groups: while second and third cohort signers generally produced universal forms, the first-cohort signers most often produced an existential quantifier (MANY). We return to this point in the General Discussion. Critically, however, even in the first cohort, universal quantifiers were not completely absent with three of the five signers producing at least one.

We also observed an emerging differentiation in the classical quantifiers for smaller and larger proportions. Signers from all three cohorts used MANY more frequently with larger proportions compared to smaller proportions and LITTLE/FEW more frequently with smaller proportions compared to larger proportions. While older signers had a strong preference for a single existential form (MANY), younger signers were more likely to produce a variety of existential quantifiers including, HALF and SOME.

Given this pattern, we might ask whether the differences we observe across cohorts reflect a refining of quantifier categories as the language develops. In other words, perhaps the initial quantificational space is carved onto classes present in the first cohort with subsequent cohorts adding novel classes, cutting the space into finer slices. Specifically, the first cohort may have two main quantifiers, LITTLE/FEW and MANY which encompasses any contexts/meanings not described by LITTLE/FEW (many, some, half, and perhaps even all). If this were the case, we would expect that ALL would be used only when contrasting with MANY. With the caveat that there was a small number of trials, we note that this was not the case. The first cohort of signers produced ALL across different contrasting contexts in this experiment. Thus, the difference we see between the cohorts in the use of ALL appears to be driven by the underuse of ALL in the existential contexts rather than a refining of an initial space of MANY to MANY and ALL.

In addition, there appears to be relative stability in how the ‘Most’ contexts are described as MANY remains the most frequent classical quantifier used by all three cohorts. But, while the ALL and MANY categories do not seem to have been drastically refined as the language developed, we do see a clear change in the terms used to describe the ‘Some’ contexts with some younger signers increasingly using other terms (e.g., HALF, SOME) in addition to LITTLE/FEW whereas older signers predominantly produce descriptions with LITTLE/FEW (see Fig. 12). Thus, this particular area of the quantifier space appears to be undergoing change as the language develops.

6. General discussion

The presence of these classical quantifiers in the first three cohorts of NSL signers supports the hypothesis that they are an early emerging property of language, perhaps because they reflect universal properties of the human mind. Indeed, our observational study suggests that these quantifiers emerged in the very first generation of signers. Below, we discuss both the aspects of quantificational language that appear to be stable across the cohorts and the aspects of quantificational language that change across these groups.

6.1. Use of existential quantifiers in universal settings

Curiously, while we did not observe stark differences across the cohorts in the presence vs. absence of quantificational language, signers from the first cohort often used existential quantifiers in settings where universal quantifiers were expected and are typically used by the later cohorts. These uses are not strictly speaking wrong. Whenever all of the bears went swimming, it is also the case that some of the bears went swimming. But, typically, language users, even children, do not generally produce an existential description in a universal context because a stronger, more informative description is possible (Grice, 1957; 1975). One possible explanation for why NSL users are producing these descriptions is because they fail to retrieve the more informative universal quantifier in a timely fashion. This could occur either because of processing limitations or because they do not know that these terms are lexical alternatives. Similar explanations have been proposed for children's failure to calculate scalar implicatures during language comprehension (e.g., Gualmini, Crain, Meroni, Chierchia, & Guasti, 2001; Pouscoulous et al., 2007; Reinhart, 1999).

A second possibility is that the first cohort signers may overuse the existential quantifiers because they have analyzed them as something simpler: functions on a single set (like numbers) rather than functions relating two sets (like classical quantifiers). Under this analysis, the existential forms are really inexact quantity expressions ("Some of the bears swim" means something like "there are some bears, who swim"). Although we cannot rule this hypothesis out, it does not provide a parsimonious explanation for the full data pattern. While the existential forms might be analyzed as functions on a single set, the negative and universal quantifiers cannot be interpreted this way ("All the girls are fishing" cannot be interpreted as "There are all girls, who are fishing"). Four of the five first-cohort signers produced at least one negative quantifier and three of the five produced at least one universal quantifier. Thus, most first cohort signers have the capacity to use quantifiers that are functions on two sets, and so it is unclear why they would avoid these terms on a subset of the trials.

Finally, the use of existential forms in universal contexts could reflect pragmatic differences in the language of the first cohort. One possibility is that the structure of lexical scales is different across the cohorts. If "some" does not form a lexical scale with "every," implicatures like "some but not all" might not exist or be weakened. Future work could investigate other scales (modified numerals, verbs) and see if similar pragmatic differences are observed. Another possibility is that older NSL signers may not realize that it is critical to be maximally informative in the universal contexts, at least in this particular task. Elicitation studies often use referential communication tasks, in which one participant attempts to describe a target item for another participant who has to find it in a set of alternatives. We chose not to do this because we suspected that this task would push participants toward a strategy of using exact numbers to maximize the likelihood of the comprehender choosing the correct picture. Future studies could use such a task, or attempt to increase the pragmatic pressure in the picture description tasks by using different contrast sets (e.g., a contrast between a target with 20 out of 20 and a foil with 19 out of 20 might create greater pressure to use a universal quantifier).

6.2. Explanations for the origin of classical quantifiers in NSL

The central question raised by these findings is: What allowed classical quantifiers to emerge so rapidly in NSL? We discuss three possibilities in turn below.

Language borrowing: One hypothesis is that these linguistic devices were borrowed from, or modelled on, quantificational forms in other languages. We consider this unlikely. The form of the quantifiers in NSL does not resemble that of quantifiers in Spanish. The individual lexical items are very different, and the quantifier does not consistently appear before the noun in NSL, as it does in Spanish. NSL has had minimal contact with other sign languages and the NSL quantifiers do not have the same form as the quantifiers in any other particular sign language, such as American Sign Language (ASL).

Backwards transmission of classical quantifiers: Another hypothesis for the emergence of quantificational language is that these classical quantifiers were created by signers in the second cohort and later adopted by first-cohort signers. The data from Study 1 suggest that this is unlikely. We observed instances of universal, existential, and negative quantifiers in the language of the first cohort signers in the 1990s when the second cohort of signers were still children. Based on findings from Pyers and Senghas (2009), we think it is unlikely that the first cohort learned and used quantifier signs from the second cohort during this time period.

The one documented case of transmission from younger signers to older signers is the case of mental state verbs. Pyers and Senghas (2009) tested the first two cohorts of NSL signers on mental state vocabulary and false-belief reasoning at two timepoints, 2001 and 2003. They found that while second cohort signers produced mental state verbs at both time points, 4 of the 8 first-cohort signers they tested in 2001 did not produce any mental state words at all. When Pyers and Senghas tested the same first-cohort signers in 2003, they found that every signer produced at least one mental state word. We do not know exactly when mental state verbs were created by the second cohort, but the pattern of findings reported in Pyers and Senghas (2009) suggest that the first cohort did not learn mental state vocabulary from the second cohort until after the second-cohort signers had entered adulthood. Specifically, between 2001 and 2003 many of the second cohort signers (M_{age} in 2001: 17; M_{age} in 2003: 19) graduated from high school and began socializing at the deaf association, increasing their interaction with first-cohort adults. We have no reason to suspect that the transmission of quantifier signs would be any different. At the time when the corpus data from Study 1 was collected, the second cohort signers were prepubescent children and the first cohort signers were in their late teens and early twenties. There is no evidence that the young adults paid much attention to the children, particularly given the findings of Pyers and Senghas (2009). Thus the appearance of classical quantifiers in historical data from the 1990s strongly suggests that these forms were created by the first generation of NSL signers.

Another possibility is that the first cohort of signers may have created some quantifier signs but then the quantifier terms evolved over time. There are two possible ways the signs may have evolved: the form may have evolved or the meaning/context in which the signs are used may have evolved.

If the forms had evolved, we should expect to see differences in the language of the first in Study 2 when compared with Study 1. In Study 1, we observed six quantifier signs. In Study 2, we observed eight quantifier signs, which included the six that we had observed in Study 1. The form of the signs we observed in Study 1 are the same as the form of the signs we observed in Study 2. In other words, the first-cohort signers produced signs that had the same form in 2015 as they did in the 90s, and the younger signers produced those same signs (rather than coming up with new signs that were then adopted by the first cohort). While we cannot say anything about the two signs we observed in Study 2 but not Study 1, we note that those two signs (HALF and one form of NONE) accounted for only a small part of our dataset from Study 2.

As for the meaning of these signs, given the limited information in the corpus, it is true that we cannot be sure that the meanings of the forms in the corpus have the same meaning as the forms in our study. We note that the forms did appear to have a similar enough meaning in the corpus that we were able to make sense of them in the context. In addition, we observed cohort differences in the use of universal quantifiers where first-cohort signers used more existential forms in universal contexts whereas second- and third cohort signers did not. Thus, it appears that one change in the context in which universal terms are used by younger signers was not adopted by the older signers at the time of Study 2.

Classical quantifiers emerge rapidly as a product of the human mind in a social context: Our findings provide strong constraints on any theory about the origins of quantifiers. Systematic means of encoding existential, universal and negative force emerged, spread and stabilized within about 40 years of the formation of this language community. Signs with apparent quantificational meanings appeared within around 15 years of the community forming. By the time we tested the participants (~25 years later) these terms had spread and stabilized and had systematically different meanings. Thus our findings clearly rule out the possibility that the creation of quantifiers is a prolonged historical process parallel to mathematics, the wheel, or the development of narrative forms (see also section 6.3). Instead, our results demonstrate that quantificational language can emerge rapidly in a new language. Our findings, however, also show that the use of quantifiers became more differentiated and robust with the second and third cohorts.

6.3. What role might cultural evolution play?

This rapid emergence of quantifiers in NSL, indicates that the linguistic and conceptual pre-requisites for quantificational language are, in a broad sense, readily available in the human mind. This does not rule out the possibility that prior cultural evolution also played a role. Our participants are immersed in society (urban Nicaragua) that uses quantificational language. While they have minimal access to the language, the concepts may permeate the culture in other ways, making them more salient and more necessary. But utility alone cannot explain how a mental ability arises: the relevant cognitive building blocks must be present and accessible. When those abilities are inaccessible, even the most important skills can fail to develop. For example, despite immersion in a numerate-rich culture and holding jobs that involve money, homesigners in Nicaragua do not have signs for exact numbers (Spaepen et al., 2011).

Any discussion of cultural evolution also requires careful thought about the range of processes that this term covers and the time scales over which they can happen. Any process of language creation involves moments in which novel forms with novel meanings are used, instances in which they are comprehended by the interlocutor, learning that occurs on the basis of that use and comprehension, and adoption and convergence in forms across the community. Depending on the size of the community and the transparency of the intended meaning, these processes might happen in minutes, months, decades, or millennia. In some cases, these processes occur with no conscious awareness or apparent effort. In other cases, they are the result of focus and effort. For example, mathematicians did not arrive at their insights implicitly, via some kind of reanalysis, but rather did so quite deliberately.

For now, it may be helpful to reserve the term cultural evolution for processes that occur on a very long time scale, and to use the term *repeated learning* for changes that occur on shorter time scales. Repeated learning within the first cohort clearly must have played some role in how different people come to use the same signs for quantificational concepts. Repeated learning, within and across cohorts, may well have played a role in how the second and third cohorts became more adept at distinguishing between universal and existential contexts and differentiating existential forms for smaller and larger proportions.

6.3. Limitations of the present study

One limitation of the present study is the necessarily small sample size. Critically, power analyses demonstrated that we had sufficient power to detect cohort differences of the size we night expect on the basis of previous studies. In fact, even with this small sample, we did detect several differences between the cohorts, including a difference in how existential forms were used in the universal context and differences in which existential forms were used for smaller and larger proportions within the existential contexts. Our most important findings, however, were those that were true across all cohorts: the frequent use of classical quantifiers and the near exclusive pairing of negative forms with negative contexts and universal forms with universal contexts. These findings rest on a larger sample of participants.

A second limitation, more critical, is that we have examined only a single emerging language, developing under particular social and educational circumstances. One case study cannot elucidate the full range of trajectories by which quantificational language may emerge. Thus future work on quantificational languages in other emerging sign languages is needed to understand whether the patterns we observed are common across contexts.

Finally, our data are ambiguous with respect to the role that communities and communication play in the creation of quantificational language. On the one hand, classical quantifiers could be concepts that are so readily available to the human mind that they quickly emerge in the communication systems of isolated individuals with no language models, such as homesigners. Alternatively, classical quantifiers could be constructed only when a community of people begin communicating together. Exact number words are clearly a case of the second kind: these words emerged in the first generation of NSL signers but they are absent in homesigners (who also do not pass conceptual tasks of exact numerosity). Examining the use of quantificational language (and concepts) in homesigners, will be critical to understanding the degree to which classical quantifiers also require input and interaction from a language community.

7. Conclusion

Across two studies, one an observational study of some of the earliest historical data available to researchers and the other an experimental study, we observed classical quantifiers in the first three age cohorts of NSL, which appear to have emerged in the very first generation of signers. These results suggest that the necessary ingredients for the creation of quantificational language are readily available in the human mind in a social context and do not require an outside linguistic model.

Author Contributions

Acknowledgments

Supplementary Materials

Due to our small sample sizes and largely categorical response patterns, not all effect sizes could be calculated for the mixed effects logistic regression models. In order to calculate effect sizes, we did the following: first, for each classical quantifier type (existential, negative, and universal), we compared its use across all conditions using the Kruskal-Wallis nonparametric test. Second, we compared the use of each quantifier type across cohorts, also using the Kruskal-Wallis test. Third, we compared the use of each quantifier type in the condition intended to elicit that type (e.g., negative quantifiers on the negative trials) by cohort using the Kruskal-Wallis test. The results of these tests are reported in the tables below. In sum, we find the same general patterns with nonparametric ordinal tests as we observed in the mixed effect logistic regression model comparisons.

Dependent variable	Independent variable	χ^2	df	p-value	ε^2
Use of Existential Quantifiers	Condition (All trials)	26.19	2	<.001	0.524
	Cohort (All trials)	4.78	2	.092	0.096

	Cohort (Existential trials)	3.79	2	.150	

Table S1. Kruskal-Wallis tests comparing use of existential quantifiers across condition and cohort.

Dependent variable	Independent variable	χ^2	df	p-value	ε^2
Use of Negative Quantifiers	Condition (All trials)	41.83	2	<.001	0.837
	Cohort (All trials)	0.44	2	.804	
	Cohort (Negative trials)	0.50	2	.779	

Table S2. Kruskal-Wallis tests comparing use of negative quantifiers across condition and cohort.

Dependent variable	Independent variable	χ^2	df	p-value	ε^2
Use of Universal Quantifiers	Condition (All trials)	28.16	2	<.001	0.563
	Cohort (All trials)	2.02	2	.364	
	Cohort (Universal trials)	5.918	2	.052	0.370

Table S3. Kruskal-Wallis tests comparing use of universal quantifiers across condition and cohort.

Across all trials, we observe an effect of condition, and a marginal effect of cohort, where the quantificational force of the produced quantifier (existential, negative, or universal) is predicted by the quantifier context (condition). Because we observed a significant interaction between condition and cohort in the mixed effect models for the use of existential classical quantifiers, we ran a Kruskal-Wallis test looking at the use of existential quantifiers on the Universal trials (see Fig. 10). In line with the results from the mixed effect models, there was a significant effect of cohort on the use of existential quantifiers in the universal trials (2(2) = 8.33, p=.016, r=.832, CI: 0.674, 0.866). Pairwise Mann-Whitney signed rank tests, corrected for multiple comparisons, revealed a significant difference between Cohorts 1 and 2 (p=.016), a significant difference between Cohorts 1 and 3 (p=.035), and a marginally significant difference between Cohorts 2 and 3 (p=.073).

References

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390-412.

Bach, E., Jelinek, E., Kratzer, A., & Partee, B. (1995). *Quantification in natural languages*. Dordrecht: Kluwer Academic Publishers.

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. In *Philosophy, language, and artificial intelligence* (pp. 241-301). Springer, Dordrecht.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *R package version*, *1*(7), 1-23.

Bihorel, S., & Baudin, M. (2018). Package 'neldermead." Available online at https://cran.r-project.org/web/packages/neldermead/neldermead.pdf

Brooks, P.J., & Braine, M.DS. (1996). What do children know about the universal quantifiers all and each? *Cognition*, *60*, 235-268.

Chomsky, N. (1988). Generative grammar. *Studies in English Linguistics and Literature*.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, *31*(5), 489-509.

Davis, H. (2010). Salish languages lack generalized quantifiers after all! Paper given at SALT 20, Vancouver, B.C.

Flaherty, M., & Senghas, A. (2011). Numerosity and number signs in deaf Nicaraguan adults. *Cognition*, *121*, 427–436.

Frege, G. (1879). Begriffsschrift, a formula language, modeled upon that of arithmetic, for pure thought. *From Frege to Gödel: A source book in mathematical logic*, *1879-1931* (pp. 1-82).

Gil, D. (1993). Nominal and verbal quantification. *STUF-Language Typology and Universals*, *46*(1-4), 275-317.

Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, *306*, 496–499.

Gualmini, A., Crain, S., Meroni, L., Chierchia, G., & Guasti, M. T. (2001). At the semantics/pragmatics interface in child language. In *Semantics and Linguistic Theory*, *Vol. 11* (pp. 231-247).

Horn, L.R. (1989). *A natural history of negation*.

Horn, L.R. (2012). Histoire d'*O: Lexical pragmatics and the geometry of opposition. In J.Y.

Béziau and G. Payette (Eds.), *The Square of Opposition: A General Framework for Cognition*, pp. 383–416, Bern: Peter Lang.

Huang, Y. T., & Snedeker, J. (2009a). Semantic meaning and pragmatic interpretation in 5-year-olds: Evidence from real-time spoken language comprehension. *Developmental Psychology*, *45*(6), 1723.

Huang, Y. T., & Snedeker, J. (2009b). Online interpretation of scalar quantifiers: Insight into the semantics–pragmatics interface. *Cognitive Psychology*, *58*(3), 376-415.

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

Huntley-Fenner, G., Carey, S., & Solimando, A. (2002). Objects are individuals but stuff doesn't count: Perceived rigidity and cohesiveness influence infants' representations of small groups of discrete entities. *Cognition*, *85*(3), 203-221.

Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child: Classification and seriation*. London, UK: Routledge and Kegan Paul.

Ionin, T., & Matushansky, O. (2006). The composition of complex cardinals. *Journal of Semantics*, *23*(4), 315-360.

Katseff, S., & Senghas, A. (2004). Effects of acquisition on the Nicaraguan Sign Language number lexicon. Delivered at *the twenty-ninth annual Boston University conference on language development (BUCLD29)*, Boston, MA.

Keenan, E. L., & Paperno, D. (2017). Overview. *Handbook of Quantifiers in Natural Language: Volume II* (pp. 995-1004). Cham, Switzerland: Springer.

Kegl, J., & Iwata, G. (1989). Lenguaje de Signos Nicaraguense: A pidgin sheds light on the “creole?” ASL. In R. Carlson, S. DeLancey, S. Gildea, D. Payne, & A. Saxena (Eds.), *Proceedings of the fourth annual meeting of the pacific linguistics conference* (pp. 266–294). Eugene, OR: University of Oregon.

Kimmelman, V. (2017). Quantifiers in Russian Sign Language. In *Handbook of Quantifiers in Natural Language: Volume II* (pp. 803-855). Cham, Switzerland: Springer.

Kocab, Senghas, & Snedeker (2016). The emergence of temporal language in Nicaraguan Sign Language. *Cognition, 156,* 147-163.

Lausberg, H., & Sloetjes, H. (2009). Coding gestural behavior with the NEUROGES-ELAN system. *Behavior Research Methods*, *41*(3), 841-849.

Matthewson, L. (2014). The measurement of semantic complexity: How to get by if your language lacks generalized quantifiers. In F. Newmeyer & L. Preston (Eds.), *Formal Complexity* (pp. 241–263). Oxford University Press.

Montague, R. (1973). The proper treatment of quantification in ordinary English. In *Approaches to natural language* (pp. 221-242). Dordrecht, Netherlands: Springer.

Nash, J.C., Varadhan, R., & Grothendieck, G. (2018). Package ‘optimx.” Available online at https://cran.r-project.org/web/packages/optimx/optimx.pdf

Partee, B. (1995). Quantificational structures and compositionality. In E. Bach, E. Jelinek, A. Kratzer, & B. H. Partee (Eds.) *Quantification in natural languages* (pp. 541–601). Dordrecht, Netherlands: Kluwer Academic Publishers.

Petronio, K. (1995). Bare noun phrases, verbs and quantification in ASL. In *Quantification in natural languages* (pp. 603-618). Dordrecht, Netherlands: Springer.

Piaget, J. (1968). Quantification, conservation, and nativism. *Science*, *162*(3857), 976-979.

Pinker, S., & Bloom, P. (1990). Natural language and natural selection. *Behavioral and Brain Sciences*, *13*(4), 707-727.

Polich, L. (2005). *The emergence of the Deaf community in Nicaraugua: With sign language you can learn so much.* Washington, D.C.: Gallaudet University Press.

Pouscoulous, N., Noveck, I. A., Politzer, G., & Bastide, A. (2007). A developmental investigation of processing costs in implicature production. *Language Acquisition*, *14*(4), 347-375.

Reinhart, T. (1999). The processing cost of reference-set computation: Guess patterns in acquisition. *UiL OTS Working Papers in Linguistics*.

Senghas, A. (1995). *Children's contribution to the birth of Nicaraguan Sign Language*. Doctoral dissertation, Massachusetts Institute of Technology.

Senghas, A., & Coppola, M. (2001). Children creating language: How Nicaraguan Sign Language acquired a spatial grammar. *Psychological Science*, *12*(4), 323-328.

Spaepen, E., Coppola, M., Spelke, E. S., Carey, S. E., & Goldin-Meadow, S. (2011). Number without a language model. *Proceedings of the National Academy of Sciences*, *108*(8), 3163-3168.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: a professional framework for multimodality research. In *5th International Conference on Language Resources and Evaluation (LREC 2006)* (pp. 1556-1559).

Wynn, K. (1996). Infants' individuation and enumeration of actions. *Psychological Science*, *7*(3), 164-169.