Cross-linguistic Patterns in the Acquisition of Quantifiers

Napoleon Katsos[a], Chris Cummins[b], Maria-José Ezeizabarrena[c], Anna Gavarró[d], Jelena Kuvač Kraljević[e], Gordana Hrzica[e], Kleanthes K. Grohmann[f], Athina Skordi[g], Kristine Jensen de López[h], Lone Sundahl[h], Angeliek van Hout[i], Bart Hollebrandse[i], Jessica Overweg[i], Myrthe Faber[ij], Margreet van Koert[k], Nafsika Smith[l], Maigi Vija[m], Sirli Zupping[m], Sari Kunnari[n], Tiffany Morisseau[o], Manana Rusieshvili[p], Kazuko Yatsushiro[q], Anja Fengler[r], Spyridoula Varlokosta[s], Katerina Konstantzou[s], Shira Farby[t], Maria Teresa Guasti[u], Mirta Vernice[u], Reiko Okabe[v], Miwa Isobe[w], Peter Crosthwaite[x], Yoonjee Hong[y], Ingrida Balčiūnienė[z], Yanti Marina Ahmad Nizar[g], Helen Grech[α], Daniela Gatt[α], Win Nee Cheong[β], Arve Asbjørnsen[γ], Janne von Koss Torkildsen[δ], Ewa Haman[ε], Aneta Miękisz[ε], Natalia Gagarina[q], Julia Puzanova[ζ], Darinka Anđelković[η], Maja Savić[η], Smiljana Jošić[η], Daniela Slančová[θ], Svetlana Kapalková[ι], Tania Barberán[c], Duygu Özge[κ], Saima Hassan[λ], Yuet Hung Chan[μ], Tomoya Okubo[ν], Heather van der Lely[ξ] †, Uli Sauerland[q], and Ira Noveck[o]*

* [a] Dept. of Theoretical and Applied Linguistics, University of Cambridge, Cambridge, CB3 9DA, UK. [b] Dept. of Linguistics and English Language, University of Edinburgh, EH8 9AD, UK. [c] Dept. of Linguistics and Basque Studies, University of the Basque Country, ES-01006 Vitoria-Gasteiz, Spain. [d] Dept. of Catalan Philology, Universitat Autònoma de Barcelona, 08913 Bellaterra, Spain. [e] Dept. of Speech and Language Pathology, University of Zagreb, Borongajska cesta 83f, Zagreb, Croatia.[f] Dept. of English Studies, University of Cyprus, 1678 Nicosia, Cyprus. [g] Independent researcher. [h] Center for Developmental & Applied Psychological Science (CeDAPS), Aalborg University, DK 9220 Aalborg East, Denmark. [i] Center for Language and Cognition Groningen, University of Groningen, Postbus 716, 9700 AS Groningen, The Netherlands. [j] Dept. of Psychology, University of Notre Dame, IN 46556, USA. [k] Amsterdam Center for Language and Communication, University of Amsterdam, 1012 VT Amsterdam, The Netherlands. [l] Hertfordshire Community NHS Trust, Welwyn Garden City, AL7 1BW, UK. [m] Institute of Estonian and General Linguistics, University of Tartu, Tartu 50090, Estonia. [n] Faculty of Humanities/Logopedics, 90014 University of Oulu, Finland. [o] L2C2, CNRS-Université de Lyon, 69675 Bron Cedex, France. [p] Department of English Philology, Ivane Javakhishvili Tbilisi State University, 0179 Tbilisi, Georgia. [q] Centre for General Linguistics, D-10117 Berlin, Germany. [r] University of Leipzig, D-04109 Leipzig, Germany. [s] Dept. of Linguistics, University of Athens, 15784 Ilissia, Greece. [t] Bar Ilan University, Ramat-Gan 52900, Israel. [u] University of Milano-Bicocca, 20126 Milano, Italy. [v] College of Law, Nihon University, Tokyo 101-8375, Japan. [w] Training Center for Foreign Languages and Diction, Tokyo University of the Arts, 110-8714 Tokyo, Japan. [x] Centre for Applied English Studies, University of Hong Kong, Hong Kong SAR. [y] Department of Second Language Acquisition, University of Maryland, College Park, MD 20742, USA. [z] Faculty of Humanities, Vytautas Magnus University, LT-44244 Kaunas, Lithuania. [α] Dept. of Communication Therapy, University of Malta, Msida MSD 2080, Malta. [β] Dept. of Psychology, HELP University, 50490 KL, Malaysia. [γ] Bergen Cognition and Learning Group, University of Bergen, 5009 Bergen, Norway. [δ]Faculty of Educational Sciences, University of Oslo, 0318 Oslo, Norway. [ε] Faculty of Psychology, University of Warsaw, 00-183 Warsaw, Poland. [ζ] Herzen State Pedagogical University of Russia, 191186, St. Petersburg, Russia. [η] Laboratory for Experimental Psychology, University of Belgrade, Studentski trg 1, Belgrade, Serbia. [θ] Institute of Slovak Studies, General Linguistics and Media Studies, Prešov University, 080 01 Prešov, Slovakia. [ι] Dept. of Speech Therapy, Comenius University, 818 06 Bratislava 16, Slovakia. [κ] Dept. of Psychology, Koç University, Rumelifeneri Yolu, 34450, Sarıyer-Istanbul, Turkey. [λ] International Islamic University/National University of Modern Languages, Islamabad, Pakistan. [μ] Department of Linguistics and Translation, City University of Hong Kong, China. [ν] National Center for University Entrance Examinations, Tokyo, Japan. [ξ] Dept. of Psychology, Harvard University, Cambridge, MA 02138.† Deceased 17 February 2014.

Submitted to Proceedings of the National Academy of Sciences of the United States of America

Learners of most languages are faced with the task of acquiring words to talk about number and quantity. Much is known about the order of acquisition of number words as well as the cognitive and perceptual systems and cultural practices that shape it. Substantially less is known about the acquisition of quantifiers. Here we consider the extent to which systems and practices that support number word acquisition can be applied to quantifier acquisition and conclude that the two domains are largely distinct in this respect. Consequently, we hypothesize that the acquisition of quantifiers is constrained by a set of factors related to each quantifier's specific meaning. We investigate competence with the expressions for 'all', 'none', 'some', 'some…not' and 'most' in 31 languages, representing 11 language types, by testing 768 5-year-old children and 536 adults. We found a cross-linguistically similar order of acquisition of quantifiers, explicable in terms of four factors relating to their meaning and use. In addition, exploratory analyses reveal that language- and learner-specific factors, such as negative concord and gender, are significant predictors of variation.

language acquisition | universals | quantifiers | semantics | pragmatics

1. Introduction

Number words and quantifiers are abstract words that denote properties of sets rather than individuals. Two-ness and all-ness in 'two/all of the black cats in the street' are not true of any individual cat, while black-ness and cat-ness are. Children display knowledge of number words and quantifiers around their second birthday, comparatively long after they have acquired concrete nouns (1, 2). As far as number words are concerned, a range of cognitive and perceptual systems support their acquisition. These include an object-tracking system, which enables the precise representation of small quantities, and an analogue magnitude system, which enables imprecise and approximate comparisons (1), as well as general principles of word-learning (3). The role of language in the acquisition of number is manifold: it can be viewed as a system of labels for expressing numerical concepts (4), a system which allows the combination of information from diverse sources (5),

Significance

While much research has been devoted to the acquisition of number words, relatively little is known about the acquisition of other expressions of quantity. We propose that the order of acquisition of quantifiers is related to features inherent to the meaning of each term. Four specific dimensions of the meaning and use of quantifiers are found to capture robust similarities in the order of acquisition of quantifiers in similar ways across 31 languages, representing 11 language types.

Reserved for Publication Footnotes

Table 1. For all quantifiers, *N* languages and types where children's performance with true and false statements was numerically higher ('>>').

Languages, out of 31				
	'All' >>	'Some' >>	'None' >>	'Some…not' >>
'All'	--	5	2	1
'Some'	26	--	14	3
'None'	29	15	--	2
'Some…not'	30	28	29	--
Language types, out of 11				
	'All' >>	'Some' >>	'None' >>	'Some…not' >>
'All'	--	3	1	0
'Some'	7	--	5	1
'None'	10	5	--	1
'Some…not'	10	10	9	--

Table 2. *N* languages and types where children rejected false statements more often than underinformative (*UI*) ones.

Languages, out of 31				
	'Some'	'Some…not'	'Most'	All three
False >> UI	31	25	24	31
Language types, out of 11				
	'Some'	'Some…not'	'Most'	All three
False >> UI	11	8	10	11

or as a provider of cues for acquisition (6, 7, 8). For example, children learning languages that distinguish between singular and plural or between singular, dual and plural morphology learn the meaning of 'one' and 'two' respectively earlier than children learning languages that do not (see 9, 10). There are also cultural practices such as the verbal count list, the recital of number words in a fixed order, 'one, two, three, …' as well as finger- or other body-part-counting routines which are widely practiced across many languages (11, 12). These systems and practices converge towards a universal order of acquisition, starting with 'one' and proceeding in line with increasing cardinality. The order itself is stable and not affected by differences between languages as regards the specific timing of the acquisition of each number word (9, 10, 13).

Quantifiers (e.g. 'none, 'some', 'all') too are properties of (or relations between) sets. The onset of the acquisition of quantifiers coincides with the acquisition of number words and some systems are likely to be implicated in the acquisition of both, e.g. principles of word learning and the role of language as a system of labels among others (3). But what about the order of acquisition of quantifiers? Is it fixed, like that of number words, or does it vary? And which systems constrain it? The perceptual object-tracking system that supports the acquisition of numbers is largely neutral to the order of acquisition of quantifiers. A set of five and a set of ten individual objects could both be referred to as 'some', 'most' or 'all' in different contexts. Moreover, there is no known routinized practice for quantifiers, such as the verbal count line or body-part counting for numbers. Even if there were to be a 'verbal quantifier line', which quantifiers would it include, and in which order? The choice is not trivial (e.g. consider 'none', 'many', 'not all', 'fewer than half') and there are multiple intuitively plausible orderings. If we were to suppose that, just as numbers are acquired in order of increasing cardinality, quantifiers are learned as a function of their increased proportion of overlap between two sets, we would predict that 'a few' and 'some' would be acquired from a very early age, and 'most' and 'all' last. Yet the evidence from corpora (14) and experiments (15, 16) reveals that, while many two-year-olds have acquired 'all', even some 7-year-old children are not fully competent with 'most'.

Overall, a simple parallelism between the order of acquisition of numbers and that of quantifiers is not fruitful and, further, does not make sense of the available evidence. While the acquisition of number words and quantifiers is supported by some shared systems, there are constraints in the order of acquisition of numbers that are not as relevant for quantifiers (such as a verbal routine). Moreover, there may well be constraints in the order of acquisition of quantifiers that do not extend to numerals.

In this paper, we hypothesize that a major constraint in the order of acquisition of quantifiers comes from the meaning of each term. Unlike number words, whose meanings vary as a function of cardinality alone, the meanings of quantifiers are varied and rich. Specific features among these word meanings are likely to play a role in their acquisition. To give substance to this distinction, consider statements such as 'All/none of the students are playing football'. 'All' is a positive and monotone increasing quantifier that licenses inferences to supersets (e.g. 'All of the students are playing *a sport*') while 'none' is a negative and monotone decreasing quantifier that licenses inferences to subsets (e.g. 'None of the students are playing football *in the rain*'). We will shortly describe this distinction formally in order to argue later that it is one of the features of meaning to play a role in acquiring quantifiers in a fixed order across languages. Of course, some languages could offer specific cues to support acquisition. For instance, they may offer additional cues that a quantifier is negative by marking negation twice, once on the quantifier itself and once with a negative particle on the verb phrase, a phenomenon known as negative concord (as in French 'aucun des élèves ne jouent au football'). In what follows, we turn to aspects of quantifier meaning and use which we argue are relevant to their order of acquisition.

2. Cross-linguistic similarities and differences

Quantifiers predicate properties of members of sets. For example, the meaning of the English quantifiers 'all' and 'some' is traditionally taken to correspond to set-theoretical logical concepts (17). Under this view, the truth-conditions of many quantified sentences are given as relations between sets as in 1, where 'iff' is 'if and only if', '$\cap$' is the intersection of two sets, '$-$' is their difference, and '$\varnothing$' is the empty set.

1. (a) 'All of the As are Bs' is true iff $A \cap B = A$
(b) 'Some of the As are Bs' is true iff $A \cap B \neq \varnothing$
(c) 'None of the As are Bs' is true iff $A \cap B = \varnothing$
(d) 'Most of the As are Bs' is true iff $|A \cap B| > |A - B|$
(e) 'Some of the As are not Bs' is true iff $A - B \neq \varnothing$

Quantified sentences have systematic entailment properties. If the sentences in 1(a, b, d) are true, then it is guaranteed that for any set B' which is a superset of B, the corresponding sentence is also true (e.g. if it is true that 'all/some/most of the students are playing football' then it is guaranteed that 'all/some/most of the students are playing a sport'). Quantifiers that guarantee inferences from sets to supersets in this way are known as *monotone increasing*. Conversely, if the sentences in 1(c, d) are true, then it is guaranteed that for any set B' which is a subset of B, the corresponding sentence is also true. Quantifiers with this property are *monotone decreasing*.

Typological research in semantics suggests that many human languages contain these and other quantifiers, and that the entailment properties of these quantifiers exhibit similarities (18). These similarities extend to considerations of quantifier usage, such as the need to be informative. For instance, speakers should not describe a situation in which all students are playing football by saying 'some students are playing football'. Under the definition in 1(b) this would be strictly-speaking true, but the speaker would be underinformative and would be potentially inviting the listener to draw further conversational inferences. These word-choices rely on norms of human rational behavior (19) and cost-benefit optimization in information exchange (20, 21). The existence of such norms is widely reported in the world's languages (22; though see 23).

Language-specific factors are also evident among quantifiers (see contributions in 24). In the following section we specify four developmental patterns that follow from cross-linguistic similarities. We then outline some of the language-specific factors that may affect acquisition. We focus on the set of four quantifiers that are the English-equivalents of 'all', 'some', 'some…not' and 'none'. These quantifiers are the basis of Aristotle's theory of syllogisms and they have held a special status in Western thought for more than two millennia (25). For reasons mentioned below, we also include 'most'.

3. Developmental generalizations

While focusing on single languages, previous studies in the processing of quantifiers (e.g. 14-16, 26, 27 a.o.) have made several generalizations that could be expected to have cross-linguistic relevance for the order of acquisition of quantifiers. Here we hypothesize that these generalizations have the status of cross-linguistically applicable constraints (see also *Discussion*). Constraint 1 concerns *monotonicity*, which we defined above. According to this constraint, children will be more successful at comprehending monotone increasing compared to monotone decreasing quantifiers (26, 28, 29). For the current study, we would expect children to show greater competence with "all" compared to "none" and with "some" compared to "some are not."

Constraint 2, *totality*, is that children are more successful at acquiring quantifiers that attribute a property to all or none of the members of a set than they are at acquiring those who attribute a property to only a part of the set (30, 31). In our data-set, this constraint will facilitate the acquisition of totality quantifiers 'all' and 'none' compared to partial quantifiers 'some' and 'some…not'.

Monotonicity and totality are independent properties. They will sometimes align to render a quantifier particularly easy or difficult for children and sometimes diverge and compete. We predict that 'all', which is a monotone increasing and a totality quantifier will be the easiest of the four Aristotelian quantifiers, while 'some…not', a monotone decreasing and partiality quantifier will be the hardest. The acquisition of 'none' and 'some' is a matter of the relative strength of the two constraints. If the advantage bestowed by totality outweighs the disadvantage of monotone decreasing, 'none' will be easier than 'some' and vice versa.

Constraint 3, *complexity*, is that children are more successful at comprehending 'some' than 'most'. In order to understand 'Most of the As are Bs', children need to be able to restrict the domain of quantification to some relevant set of As in the universe of discourse and then compare the cardinalities of the set of As that are Bs to the set of As that are not Bs (see also 32). However, 'Some As are Bs' is simpler because in this case children do not need to restrict the quantifier to a specific set of entities or to compare cardinalities. They can simply treat 'Some students like football' as logically equivalent to 'There is at least one entity that is both a student and likes football' (33).

Finally, Constraint 4, *informativeness*, is that children will be stricter towards violations of truth than towards violations of pragmatic felicity. That is, children do not reject utterances that are underinformative (e.g. saying 'some' when 'all' is true) to the same extent as utterances that violate truth (e.g. saying 'some' when 'none' is true) nor to the same extent as adults (27, 32, 33). We therefore expect that children will accept underinformative utterances more often than false ones regardless of the language they speak. In our data-set, this means that children are more likely to reject a false statement with 'some', 'some…not' and 'most' than an underinformative one (and at rates that are distinguishable from adults).

These predictions are summarized in 2(a-c) below, ('>>' implies higher performance, and '/' no prediction):

2. (a) Constraints 1 & 2: 'all' >> 'none' / 'some' >> 'some…not'

(b) Constraint 3: 'some' >> 'most'

(c) Constraint 4: False >> underinformative for 'some', 'some…not' and 'most'

In addition to these four factors that may affect the acquisition of quantification in similar ways across languages, language-specific properties may have an important role too. The explicit presence of a partitive marker (such as 'of the' in English) may positively affect children's performance with underinformative utterances (27) by drawing attention to the divisibility of the reference set. Syntactically, negative concord may be a significant predictor, with the presence of two negative markers highlighting the fact that the utterance contains a negative quantifier. Finally, a range of non-linguistic factors may also be important predictors of children's performance. These include biological factors such as gender and age, and social factors such as socio-economic and educational status (e.g. whether children are enrolled in formal schooling at time of testing).

4. The experiment

As part of a larger project known as the COST Action A33 (see acknowledgements footnote), the empirical investigation focused on the comprehension of quantified sentences by 768 children (mean age: 5;5; age range 5;00 – 5;11; 398 of them were female) and 536 adult participants (all adults were over 18 years of age; 293 adults were female – due to experimenter error, the gender of 46 adults was not recorded). The participants spoke one of 31 languages, Basque, Cantonese (Yue) Chinese, Catalan, Croatian, Cypriot Greek, Danish, Dutch, English, Estonian, Finnish, French, Georgian, German, Greek, Hebrew, Italian, Japanese, Korean, Lithuanian, Malay (Kuala Lumpur variety), Maltese, Mandarin Chinese, Norwegian, Polish, Russian, Serbian, Slovak, Spanish, Tamil, Turkish and Urdu. This sample contains representatives of fifteen language genera (Baltic, Chinese, Finnic, Germanic, Greek, Indic, Japonic, Karto-Zan, Korean, Malayo-Sumbawan, Romance, Semitic, Slavic, Southern Dravidian and Turkic). These belong to eleven language types (seven of the main language families in the world, Afro-Asiatic, Altaic, Austronesian, Dravidian, Indo-European, Kartvelian, Sino-Tibetan, and Uralic/Finno-Ugric, as well as three language-isolates, Basque, Japonic and Korean, classified according to 34). Details of the languages' properties are given in Table S1. In the main part of the task, participants were presented with five boxes and five objects. Between none to five of the objects were inside the boxes for any test item. Participants then heard a description containing one of the five quantifiers and had to judge if the description was "right" or "wrong" for the visual display. Details of the test procedure are presented in the Methods section.

4.1 Results

The results for child and adult participants per language are presented in Tables S2 and S3. Across all languages and expressions, adult responses were on average 99% correct in the true or false conditions. These ceiling adult data validate the task as a test of competence with quantification and are no longer discussed. Eighty-four per cent of adult responses to under-informative items were rejections; this less-than-perfect consistency accords with previous literature (32 among others) and is discussed in the context of Constraint 4.

Across all languages and expressions, child responses were on average 82% correct in the true or false conditions and 51% of responses in under-informative conditions were rejections. Starting with Constraint 1, monotonicity, we first report child performance with each of the monotone increasing quantifiers in the data-set, 'all' and 'some', as compared to the performance with each of the monotone decreasing quantifiers ('none' and 'some…not'). Performance with 'all' was numerically higher than with 'none' – the monotone decreasing quantifier which is matched with 'all' for totality – in 29/31 languages. The exception was Korean (we consider 'exceptions' those languages where the numerical difference was the opposite of the one expected), while there was no numerical difference in English. Turning to 'all' and 'some…not' – the monotone decreasing expression which is not matched to 'all' for totality – children performed better with 'all' in 30/31 languages, with no differences in Georgian.

In 28/31 languages children performed better with monotone increasing 'some' compared to 'some…not', the monotone decreasing quantifier which is matched for totality (Catalan was an exception, with no difference in English and Georgian). Children performed better with 'some' than with 'none' in 15/31 languages (the exceptions being Cantonese, Catalan, Dutch, English, Estonian, Finnish, French, German, Greek, Japanese, Polish, Serbian, Slovak, Turkish; no differences in Cypriot Greek and Georgian).

Overall, when keeping the setting of totality constant, that is, comparing the two totality quantifiers, 'all' and 'none', with each other and the two partiality quantifiers, 'some' to 'some…not', with each other, the monotone increasing quantifiers give rise to better performance than the corresponding monotone decreasing ones in 27/31 languages (Catalan, English, Georgian and Korean being exceptions).

Turning to totality, performance with 'all' was higher than with 'some' (which is the quantifier with the same setting of monotonicity) in 26/31 languages (with Korean, Malay, Maltese and Russian as exceptions, and no differences in Georgian). Children performed higher with 'all' than with 'some…not' (which is the quantifier with a different value for monotonicity) in 30/31 languages, with no differences in Georgian.

Performance with 'none' was higher than with 'some…not' which is matched for monotonicity in 29/31 languages (with Tamil as exception and no differences in Georgian) and higher with 'none' than with 'some', which has a different setting for monotonicity, in 14/31 languages.

Overall, when keeping monotonicity stable, totality quantifiers 'all' and 'none' give rise to better performance than the corresponding partiality ones ('some' and 'some…not' respectively) in 25/31 languages (Georgian, Korean, Malay, Maltese, Russian, Tamil being exceptions). Visual inspection of Table 1 shows that the order predicted by Constraints 1 and 2 is indeed upheld, with 'all' being the easiest quantifier for 5-year-olds across the languages in our sample, and 'some…not' the hardest. The two constraints have relatively equal weight, with no consistent order of acquisition between 'some' and 'none'.

Multivariate analyses were also performed. These revealed main effects of language, monotonicity and totality along with higher performance when the correct answer was rejection. A small effect of gender (boys outperforming girls) was also obtained, but we found no significant effect of age. See S4.

We also conducted parallel analyses using language genus (n=15) and language type (n=11; family or isolate) in place of individual languages, along with analyses without any language variable at all. These returned a significant effect of language genus and type, but in all cases, model comparison using the Akaike Information Criterion (AIC; 35) revealed that the inclusion of any one of the language variables resulted in the model being overfitted compared to a model with no language variables, hence that the inclusion of language, genus or type in the model was not statistically justified. Likewise, models positing an interaction of monotonicity or totality with the language variables were overfitted. Therefore, the data are most appropriately modeled by positing effects of monotonicity and totality but no effect of language, whether at the level of each individual language, genus or type. Put in another way, children were more successful with the acquisition of quantifiers in some languages compared to others, but the main effects on the order of acquisition that we hypothesized, monotonicity and totality, were upheld in the data-set regardless of the specific language (or language genus or type) the children were learning.

Turning to Constraint 3, the hypothesis that 'some' would be mastered earlier than 'most' on account of its semantic simplicity was borne out numerically in all 31 languages in our sample. The effect of complexity was corroborated through multivariate analyses as with Constraints 1 and 2. Model comparison indicated that models that included language, genus, or type (or an interaction of complexity by language, genus, or type) were overfitted by comparison with models that did not. A small effect of gender (boys outperforming girls) was obtained, but no significant effect of age. See S5 for details.

Finally we consider Constraint 4, underinformative uses of 'some', 'most' and 'some…not'. In comparison to the false statements with the same expression, children rejected underinformative uses less often in all 31 languages. Looking at each expression on its own, underinformative 'some' was rejected less often than false 'some' in every language. This preference held for 'some…not' in 25/31 languages (the exceptions being Croatian, Hebrew, Malay, Maltese, Mandarin, and Tamil) and for 'most' in 24/31 languages (the exceptions being Danish, English, Finnish, French, Norwegian, Polish, Slovak). See Table 2.

For Constraint 4 we also discuss the adult data, because the adults rejected underinformative statements more frequently than children did (84% compared to 51%) but they did not reach ceiling. Looking at all three quantifiers, adults rejected underinformative uses less often than false ones in 28/31 languages. Cantonese was an exception due to two erroneous responses among false statements and ceiling performance in the underinformative conditions. Russian and Urdu showed no differences, with both false and rejected underinformative conditions being at ceiling in both languages. Furthermore, Constraint 4 held in 25/31 languages for the case of 'some' (with Basque, Croatian, Cantonese, Georgian, Russian and Urdu showing no difference), in 27/31 for 'some…not' (with Cantonese as an exception and Georgian, Russian and Urdu showing no difference), and 25/31 for 'most' (with Cantonese as an exception and English, Mandarin, Russian, Turkish and Urdu showing no difference). Therefore, not only do the child data support Constraint 4, the adult data do too.

We performed multivariate analyses for each of the quantifiers 'some', 'some…not' and 'most' for the child data. In each case, highly significant main effects of language and informativeness were shown, with underinformative statements being rejected less often than false ones. No effects of gender or age were obtained. See S6. Model comparison again suggested that models including language, genus or type or their interactions with informativeness were overfitted.

The analyses for Constraints 1–4 for the child data can be supplemented by comparisons with what would be expected if performance were guided by chance. Everything else being equal, 27/31 languages accorded with monotonicity (Catalan, English, Georgian and Korean being exceptions), 25/31 with totality (Georgian, Korean, Malay, Maltese, Russian, Tamil being exceptions), and all 31 accorded with complexity and with informativeness for all quantifiers. Each of these patterns is more consistent than if the distribution was random ($p < 0.01$ by the Sign Test). See Fig.S1 and Fig. S2.

Having demonstrated our effects of interest and having further documented that there is variability between languages, we then explored whether this latter variability is explicable by other linguistic factors or features of the learners in our sample. Exploratory analyses suggest that attending formal school at the time of testing was a significant facilitating factor ($p<.001$) as were learning languages that use negative concord ($p<.001$) and learning expressions with a partitive marker in the case of 'some' ($p<.05$). As our language sample is not balanced with respect to these properties, we do not draw firm conclusions here.

5. Discussion

The descriptive reports and the statistical modeling analyses suggest that our hypothesized Constraints 1-4 are valid generalizations about the order of acquisition of quantifiers across the languages in our sample. These constraints were posited on the basis of generalizations made in previous research in single languages (e.g. 14-16, 26, 27 a.o.) and the present findings confirm their relevance to acquisition more widely. However, further research is required to elucidate their nature and produce theoretical models from which they would follow. For example, Constraint 1, monotonicity, is closely related to negation (29) in that all negative quantifiers are monotone decreasing, but not vice versa. Since both monotone decreasing expressions in our sample, 'none' and 'some…not' contain negation, further work could reveal whether the effects we obtained here are due to monotonicity, negation, or both.

As regards the exceptions in our sample, an important question is whether there was systematicity among the languages that did not conform to the hypothesized constraints. Two observations suggest this is not the case. First, no language or language type violated more than one constraint, except Georgian, which violated two. Second, in Georgian (as well as in other languages), the violations were evidenced in cases of ceiling performance.

This leads to the issue of generalizability of the patterns in other languages and for other quantifiers. Our sample consists of representatives of 11 language types. While there is an over-representation of Indo-European languages in our sample, the diversity of distinct language types in our sample is squarely within the range used for state-of-the-art comparative linguistic (e.g. 24) and psycholinguistic research (22). Of course, extrapolating from patterns observed in this sample to universal patterns should always be done with caution and as a working hypothesis only.

Similar considerations apply when extrapolating to quantifiers not tested here. For example, many languages have more than one universal quantifier, including the English-equivalent of an 'each' quantifier that is used for distributive quantification (36 reports eight different universal quantifiers in Malagasy which differ on the dimension of distributivity). The prediction is that the effects we obtained here should hold, as long as the appropriate considerations are taken into account. Turning to the case of 'each', monotonicity and totality should facilitate its acquisition across different languages but distributivity itself may be an additional important – facilitating or hindering – factor.

In terms of explaining the cross-linguistic variation, where the acquisition of quantifiers was more successful in some languages compared to others, exploratory analyses found that language-specific features, such as using negative concord and partitive markers had a facilitating effect. We hypothesize that negative concord may serve to better highlight that a quantifier is negative, and additionally highlight the contrast between negative and positive quantifiers. Partitives highlight that these expressions are related to parts of sets. Cross-linguistic variation may also be due to linguistic factors that we did not model in our analyses (e.g. agreement, the number of competing expressions and the overlap of their meaning). Clearly, further research on this topic is called for.

Exploratory analyses also revealed an effect of attending school at time of testing. We do not believe that the effect is related to explicit instruction about quantifiers, as all the teachers and caregivers of the children we recruited reported that quantifiers were not part of the curriculum or any extra-curricular activity. Instead, we hypothesize that attending school raises the children's readiness for activities of the kind that we administered. We also found that age was not a significant predictor of success. We believe that this was due to the restricted age-range which was part of the selection criteria (5;00 to 5;11).

Our analyses also found a gender effect, whereby boys in this study outperformed girls in the acquisition of the true or false meaning of the quantifiers (see S4-S5) but there were no differences when it came to informativeness (see S6). Linguistic skills are generally more advanced among girls than among boys (37, 38). An investigation of over 13,000 children in 10 European linguistic communities suggests that these advantages are robust across different languages (38), even though the level of overall linguistic attainment differed. Research on gender and mathematical competence suggests that there are wide-spread similarities between boys and girls (39). Nevertheless, a specific and small advantage is reported for boys for mathematical reasoning, perhaps reflecting higher aptitude with logical and set-theoretical concepts (39). Conversely, an advantage specific to arithmetic is reported for girls, which seems to be attributable to the girls' higher verbal skills which are implicated in arithmetical processing (40).

To the extent that these gender differences are robust, the language of quantification brings them into competition. Girls in our sample may have benefitted from an overall advantage in language skills and arithmetic and counting, while boys may have benefitted from an advantage with set-theoretical concepts with the latter being more critical for the specific task than the former. We should note that our analyses for gender effects were exploratory and that future studies should take into account several potentially confounding factors (40).

Before we conclude, we need to address an alternative interpretation of the findings. That is, perhaps the patterns obtained here reflect competence with counting and checking the objects that need to be verified as belonging to a set (rather than competence with the meaning of a quantifier). We can reject this interpretation for two reasons. First, counting and verifying sets with up to five members, the maximum required in this task, was part of the selection criteria (see Methods). Moreover, increased demands on counting and verification complexity do not make correct predictions in this data-set. To take but one example, consider 'none' and 'some…not'. When 'some…not' is true, that is, when two out of five objects are in the boxes, in a random selection checking procedure given five objects, 'some…not' requires checking the position of 1.5 objects on average against the boxes. When false, that is five out of five objects are in the boxes, 'some…not' requires checking the position of five objects. For 'none', this is five objects when 'none' is true (and five out of five objects are outside the boxes) and two objects when false (when two out of five objects are in the boxes). In sum, to give the correct response to 'some…not' in true and false conditions participants need to check 6.5 objects on average against the boxes, and for 'none' seven. If it were the case that counting and verification complexity were primarily responsible for performance, 'some…not' ought to be easier than 'none'. At the very least there ought to be no major difference. Yet 'none' is easier than 'some…not' in 29/31 languages and 9/11 types, as predicted by constraint 2, totality. Of course verification and counting are an important component of success with tasks like ours and further research could identify their role for younger children in order to determine which specific verification strategy is implemented for each quantifier (see e.g. 26, 41).

6. Conclusion

In this paper we investigated the order of acquisition of five common quantifiers and hypothesized four cross-linguistic constraints on their acquisition, based upon considerations of their meaning and use. A cross-linguistically similar order of acquisition emerged in a sample of 31 languages. This order accorded with the constraints we posited, supporting the claim that they are potential universals in the acquisition of quantification. This is in line with recent proposals favouring the existence of extensive cross-linguistic similarities in language meaning and use (22, 42). However, we also found that language-specific features, such as whether a language uses negative concord, have a significant effect on the learners' performance, as do social and biological factors.

Methods

See S2 and S3 for details of child and adult participants per language. The actual quantifiers used in each language were selected by researchers who were native speakers of that language. Where more than one lexical item was available, the choice was guided by considering which item would be most familiar to children. Where possible, this decision was informed by investigating corpora of child-directed speech; in other cases, researchers consulted colleagues and/or school-teachers. See table S7 for materials and glosses.

Children were tested at nurseries or primary schools following the ethical protocols designated by the host institutions of the participating researchers. They were administered the 'Cavegirl task' which was designed to test the comprehension of quantified sentences (16). In this task the Cavegirl is asked to say "How many toys are in the boxes" in visually presented situations. In each trial, the Cavegirl produces a single utterance of the type '[Quantifier] (of the) [objects] are (not) in the boxes'). Children are then asked to evaluate whether what the Cavegirl said was "right" or "wrong" and if they say "wrong" to justify why. Two types of visual situations are used for each quantifier tested, one which renders an utterance with this quantifier true and informative and one which renders an utterance false. For 'some', 'most' and 'some…not', there is also a third type of display that renders an utterance true but pragmatically underinformative (where all the objects are in the boxes for 'some' and for 'most' and where none of the objects are in the boxes for 'some…not').

The task is preceded by a warm-up session where children are familiarized with the Cavegirl, the task demands, and the pictures of the objects mentioned in the sentences. The first five items of the task test the comprehension of number words 'one' to 'five', to ensure that children can make correct judgments about quantity when simple counting is involved. Children that did not perform correctly with all five number words did not continue with the main task. This resulted in less than 5% of children not continuing. All justifications of rejections in the main task, whether correct or incorrect, mentioned a quantity-related word or deictic expression often combined with a spatial expression (e.g. "Because these are out"), which suggests that children responded based on the appropriateness of the quantifier rather than some other aspect of the sentence. See (16) for further details of the task administration and a full list of items in their respective visual situations as well as sample visual displays.

Acknowledgements: This research was made possible by funding from COST Action A33 'Cross-linguistically robust stages of children's linguistic performance'. In addition, NK, CC and IN were supported by the ESF Euro-XPrag network; NK, CC and NS by the UK ESRC XPrag-UK network; NK by the UK BA grant SG090676; AG by MINECO project FFI2014-56968-C4-1; KKG by a University of Cyprus Project (no. 8037-61017); KJdL and LS by the Danish FKK, grant no. 09-063957; MV and SP by the Estonian Science Foundation grant no. ETF7492 and the Estonian Research Council grant no. SF0180056s08; AA and JvKT by a grant from the L. Meltzers Høyskolefond; EH and AM by grant no. 809/N-COST/2010/0 from the Polish Ministry of Science and Higher Education and National Science Centre; KY and US by the European Commission for Education and Culture grant no. 135295-LLP-2007-UK-KA1SCR and by the BMBF grant no. 01UG0711; DA, MS, and SJ by grant ON179033 (2011-2014) by the Serbian Ministry of Education, Science and Technological Development..

1. Feigenson L, Dehaene S, Spelke ES (2004) Core systems of number. *Trends Cogn Sci* 8(7):307–314.
2. Wynn K (1992) Children's acquisition of the number words and the counting system. *Cognit Psychol* 24:220-51.
3. Sullivan J, Barner D (2011) Number words, quantifiers, and principles of word learning. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2:639-645.
4. Leslie AM, Gelman R, Gallistel CR (2008) The generative basis of natural number concepts. *Trends Cogn Sci* 12(6):213–218.
5. Spelke ES, Tsivkin S (2001) Language and number: A bilingual training study. *Cognit* 78(1):45–88.
6. Bloom P, Wynn K (1997) Linguistic cues in the acquisition of number words. *J Child Lang* 24(3):511–533.
7. Carey S (2009) *The origin of concepts* (Oxford Univ Press, New York).
8. Xanthos A et al. (2011) On the role of morphological richness in the early development of noun and verb inflection. *First Language* 31(3):461–479.
9. Almoammer A, Sullivan J, Donlan C, Marušič F, Žaucer R, O'Donnell T, Barner D (2013) Grammatical morphology as a source of early number word meanings. *Proc Natl Acad Sci USA*, 110:18448-18453.
10. Sarnecka BW, Kamenskaya VG, Yamana Y, Ogura T, Yudovina YB (2007) From grammatical number to exact numbers: early meanings of 'one', 'two', and 'three' in English, Russian, and Japanese. *Cognit Psychol* 55(2):136–168.
11. Gelman R, Gallistel CR (1978) *The child's understanding of number* (Harvard University Press, Cambridge MA).
12. Saxe GB (1981) Body parts as numerals: A developmental analysis of numeration among the Oksapmin in Papua New Guinea. *Child Development* 52:306–316.
13. Barner D, Libenson A, Cheung P, Takasaki M (2009) Cross-linguistic relations between quantifiers and numerals in language acquisition: Evidence from Japanese. *J Exp Child Psychol* 103: 421–440.
14. Hanlon C (1988) in *The development of language and language researchers: essays in honor of Roger Brown*, ed Kessel FS (Erlbaum, Hillsdale NJ), pp 65–78.
15. Barner D, Chow K, Yang SJ (2009) Finding one's meaning: A test of the relation between quantifiers and integers in language development. *Cognit Psychol* 58(2):195–219.
16. Katsos N, Andrés Roqueta C, Estevan RAC, Cummins C (2011) Are children with Specific Language Impairment competent with the pragmatics and logic of quantification? *Cognit* 119:43–57.
17. Barwise J, Cooper R (1981) Generalized quantifiers in natural language. *Linguist Philos* 4:159–219.
18. von Fintel K, Matthewson L (2008) Universals in semantics. *The Linguistic Review* 25(1-2):139–201.
19. Grice HP (1975) *Studies in the way of words* (Harvard University Press, Cambridge MA).
20. Levinson SC (2000) *Presumptive meanings* (MIT Press, Cambridge, MA).
21. Sperber D, Wilson D (1986/1995) *Relevance: communication and cognition* (Blackwell, Oxford).
22. Stivers T et al. (2009) Universals and cultural variation in turn-taking in conversation. *Proc Natl Acad Sci USA* 106(26):10587–10592.
23. Keenan, EO (1974) in *Studies in linguistic variation*, ed Fasold RW, Shuy RW (Georgetown University Press, Washington DC) pp 255–268.
24. ed Keenan EL, Paperno D (2012) *Handbook of quantifiers in natural language* (Springer, Netherlands).
25. Bonevac D (2012) in *Logic: a history of its central concepts*, ed Gabbay DM, Pelletier FJ, Woods J (Elsevier, Oxford) pp 63–126.
26. Deschamps, I, Agmon G, Loewenstein Y, Grodzinsky Y (2015) The processing of polar quantifiers and numerosity perception. *Cognit* 143:115–128.
27. Pouscoulous N, Noveck IA, Politzer G, Bastide A (2007) A developmental investigation of processing costs in implicature production. *Lang Acquis* 14:347–376.
28. Geurts B, Katsos N, Cummins C, Moons J, Noordman L (2010) Scalar quantifiers: logic, acquisition, and processing. *Lang Cogn Process* 25(1):130–148.
29. Horn LR (1989). *A natural history of negation* (University of Chicago Press, Chicago).
30. Hackl M (2009) On the grammar and processing of proportional quantifiers: Most versus more than half. *Nat Lang Semantics*, 17:63–98.
31. Keenan EL (1996) in *The handbook of contemporary semantic theory*, ed Lappin S (Blackwell, Oxford) pp 41–64.
32. Noveck IA (2001) When children are more logical than adults. *Cognit* 86: 253–282.
33. Guasti MT et al. (2005) Why children and adults sometimes (but not always) compute implicatures. *Lang Cogn Process* 20(5):667–696.
34. Dryer MS, Haspelmath M (2013) *The world atlas of language structures online* (Max Planck Institute for Evolutionary Anthropology, Leipzig). Available online at http://wals.info, accessed 2014-10-09.
35. Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* 19(6):716–723.
36. Keenan EL (2008) in *Quantification: a cross-linguistic perspective*, ed Matthewson L (John Benjamins, Amsterdam) pp 319–352.
37. Burman DD, Bitan T, Booth JR (2008) Sex differences in neural processing of language among children. *Neuropsycholog* 46:1349–1362.
38. Eriksson M et al. (2012) Differences between girls and boys in emerging language skills: evidence from 10 language communities. *Br J Dev Psychol* 30:326–343.
39. Hyde JS, Lindberg S, Linn M, Ellis A, Williams C (2008) Gender similarities characterize math performance. *Science* 321:494–495.
40. Wei W, Lu H, Zhao H, Chen C, Dong Q, Zhou X (2012) Gender differences in children's arithmetic performance are accounted for by gender differences in language abilities. *Psychol Sci* 23(3):320–330.
41. Lidz J, Pietroski P, Hunter T, Halberda J (2011) Interface transparency and the psychosemantics of 'most'. *Nat Lang Semantics* 19:227–256.
42. Levinson SC (2016) Turn-taking in human communication, origins, and implications for language processing. *Trends Cogn Sci* 20(1):6–14.