I Perspectives on mathematics and language of different disciplines

Caleb Everett

The diversity of linguistic references to quantities across the world's cultures

1 Introduction: Defining core concepts

It is challenging, perhaps impossible, to discuss "numbers" without bringing to bear particular assumptions of one's culture, language, theoretical bias, or some combination thereof. What are numbers, exactly? Are they innate concepts given to us by natural selection? Are they entities that exist in nature, awaiting discovery by the brains of humans and other animals? Are they cultural and linguistic constructs that have gradually accrued across the world's populations in different but constrained ways? Some scholars would offer affirmative answers to only one of the three preceding questions, while others might offer positive answers for all three. Volumes have been written on these possible perspectives and, perhaps, many of those volumes are of little relevance to those interested in more quotidian, and perhaps more significant, pedagogical concerns associated with numbers. Nevertheless, it is useful to have some basic agreement as to what we mean when we talk about learning numbers and the basic mathematical principles revolving around them – to have some shared understanding about what numbers even are. In this chapter I will focus on the last of the three questions above, outlining in basic form the cross-linguistic variation that exists vis-à-vis spoken number systems to illustrate how such systems have accrued in variable ways across human cultures – even if the relevant variations are constrained in some ways. The survey presented should, I hope, benefit scholars interested in mathematical pedagogy who are not entirely familiar with the extent of cross-cultural variation in the number systems of the world's languages.

Before embarking on the survey, though, allow me to establish the definitions of three terms that I will be using during its course. These definitions reflect my own theoretical predispositions, informed as they are by cross-cultural and cross-linguistic data. The three key terms and associated definitions I will employ are "quantical," "numerical," and "numbers." The definitions are grounded in other work, primarily Núñez (2017), though related terms and definitions have been presented by a variety of scholars. I begin with "numbers," which I define as verbal symbols representing precise quantities. ~~(In contrast to written numerals, which are not discussed in this contribution.)~~ Given that verbal symbols for precise quantities have primacy both ontogenetically and cross-culturally in

AU: Please rephrase the sentence "In contrast to written numerals, which..." for completeness.

https://doi.org/10.1515/9783110661941-001

our species, when contrasted to nonverbal symbols, I find it useful to interpret them as the default form of numbers. This may seem odd in cultural contexts in which written symbols are sometimes interpreted as equally (or more?) basic, learnable units, but I believe that a focus on numbers-as-words is a useful reminder of the primacy of verbal symbols for representing precise quantities. Judging from the cross-cultural data, humans' most basic symbolic tools for manipulating quantities are verbal (Everett, 2017).

The distinction between "quantical" and "numerical" concepts is more recent and esoteric, but I believe it to be extremely useful and well motivated. For a fuller discussion of the merits of this distinction, I refer the reader to Núñez (2017). The chief motivation is that much research in psychology refers to basic and native "numerical" cognition, putatively shared by our entire species, that appears to be neither basic nor native once the extent of cross-cultural diversity in quantitative cognition is surveyed with sufficient care. Nevertheless, it is generally agreed that all humans do share some basic native capacities for quantity discrimination. For instance, humans can generally distinguish small quantities (1, 2, and 3) from each other precisely without training (as can the members of some other species). Humans can also approximately discriminate larger sets of items, for instance, eight sticks from sixteen sticks, presuming that the ratio between the sets is large enough. (This ability is also phylogenetically primitive – some have suggested it stretches back to the first vertebrates.) These basic quantitative reasoning skills are not apparently contingent on cultural scaffolding, but they are not really "numerical" in that they offer no means of delimiting, for instance, five from six items with consistency. Numbers like "five" and "six" do not simply follow from our native quantitative capacities; they must be crafted and honed by distinct cultural practices that rely on those capacities. These practices allow us to transfer our modest native exact quantity recognition into the realm of larger quantities. For such reasons, it is not particularly useful (from my perspective anyhow) to refer to native quantical abilities, shared with other species, as "numerical," or to liken them to a "number sense." Terms like "number sense" may give the false impression that we are somehow born with numbers in our heads or are wired to learn basic arithmetic (Dehaene, 2011). In the words of Núñez:

> Humans and other species have biologically endowed abilities for discriminating quantities. A widely accepted view sees such abilities as an evolved capacity specific for number and arithmetic. This view, however, is based on an implicit teleological rationale, builds on inaccurate conceptions of biological evolution, downplays human data from non-industrialized cultures, overinterprets results from trained animals, and is enabled by loose terminology that facilitates teleological argumentation. (2017:409)

Given my shared desire to avoid teleological argumentations where they are not warranted, and given this chapter's focus on cross-cultural variability, I adopt the terminological distinction proffered by Núñez, the distinction between "quantical" concepts and "numerical" concepts. The former term refers to humans' native, biologically endowed capacities for differentiating quantities in generally coarse ways. The latter term, "numerical," refers to exact, symbolic practices evident when humans use "numbers." Framed differently: The existence of quantical cognition is a necessary condition for the flowering of numerical cognition, but it is, critically, not a sufficient condition. Maintaining a distinction between "quantical" and "numerical" cognition is particularly useful as a background for discussing the extent of cross-cultural variability in the ways that people talk about quantities, and the potential relevance of that extensive variability to mathematical pedagogy. It is important to dissociate the universals of human quantical thought from the cross-cultural variability of numerical thought and numbers. This clear dissociation could positively impact efforts to more effectively convey numerical concepts to individuals across the world's cultures.

So, to be clear, this contribution aims to shed light on the diversity of numbers in the world's languages in the expression of numerical concepts, and also will survey some differences in how languages describe quantical concepts. Approaches to the pedagogy of arithmetic could only benefit, I hope, from an understanding of commonalities and differences in the ways the world's languages refer to such concepts. These could offer some insights into the best ways to approach, for instance, cross-culturally effective instruction strategies. (I leave it to the experts on pedagogy, however, to judge how the findings discussed here might benefit mathematical instruction across cultures.) At the least, such commonalities and differences can hopefully give the reader a better sense of just how typical or atypical our own linguistic strategies for encoding numerical and quantical concepts are, when considered in the light of the typological data. By examining an adequately representative sample of number systems in the world's languages we can, inter alia, better understand which numerical concepts are most easily acquired by the members of our species.

2 Cross-population differences in grammatical number

The grammars of the world's languages often refer to quantical concepts, what is commonly referred to as "grammatical number." Grammatical number refers to a variety of phenomena that denote distinctions between small precise quantities and large imprecise quantities (e.g., singular vs. plural), or between small precise quantities (e.g., singular vs. dual). Grammatical number markers take many forms, including noun suffixes and prefixes, verb suffixes and prefixes, and many more. In English, for example, suffixes are added to nouns to demonstrate whether there is one or more than one of an item or entity to which the speaker is referring. In the languages in which grammatical number exists, it serves overwhelmingly to distinguish between sets of exactly one (singular) and more than one (plural). In rarer cases grammar is also used to distinguish one, from two, from more than two items. Languages with that kind of grammatical number are said to have singular, dual, and plural marking. Rarer still are languages that have singular, dual, trial, and plural marking. So grammatical number is always used to designate sets of items (1, 2, 3, or many) that humans are capable of discriminating via their native quantical cognition, as defined above.

Grammatical number refers only to small quantities precisely, and to large quantities approximately. In this way its function is limited, but in another sense its function is very robust: Languages that have grammatical number often use it to obligatorily denote the quantity of reference, and this obligatory status means that it is extremely pervasive in speech. In this chapter alone there are hundreds of cases of grammatical number inflected on verbs and nouns. English learners, whether children or adults, must learn the ways of adding regular plural markings, not to mention irregular plural markers. They must also learn that some nouns, say, "deer," are not marked at all in the plural. More broadly, they learn that the quantity of referents is always relevant, even if only in approximate ways, during communication.

This is not the case in many of the world's cultures. In fact, in about 10% of the world's languages, there is no grammatical means of designating the number of referents to which a speaker is referring. For example, the Karitiâna language, on which I have done a fair amount of research, has no nominal plurality. Consider the following phrases from that language:

(1) *myjyp ambi*
three house
"Three houses."

(2) y-ambi
1st.Singular.Possessive-house
"My houses."

(3) *ombaky* *naokyt* *taso*
jaguar killed man
"The jaguar(s) killed the man/men."

(4) *yj-pyt* *ombaky*
our-hand jaguar
"Five jaguars."

As we see in (1) and (2), the word for house does not change even when there are many houses being referred to. The same is true of "jaguar," and "man," as seen in (3) and (4), because all the nouns in the language do not denote quantity distinctions.

There are many languages like Karitiâna scattered around the world. In a survey of data of 291 languages representing many distinct language families and geographic regions, the linguist Martin Haspelmath found that about 10% (n = 28) of the languages were like Karitiâna, with no nominal plurality evident in their grammars (Haspelmath, 2013). ("Nominal plurality" refers to cases in which the quantity of an item referred to by a noun is denoted in the grammar, typically with a suffix on the noun.) In another 19% (n = 55) of the languages, nominal plurality was found to be optional in all cases. So rather than saying something like, for instance, "three cars," one could say "three cars" *or* "three car," and either would be grammatically correct. There is a sense in which this is intuitive, as the *-s* suffix in a phrase like "three cars" is, after all, redundant, encoding information about plurality that is already contained in the preceding number word. In other cases the plural marking may prove quite informative. For example, the interpretation of clause (3) could vary significantly. Did one jaguar kill one man? Did one jaguar kill many men? Did many jaguars kill many men? Did many jaguars kill one man? In actuality, though, context and real-world prior information (e.g., that jaguars are fairly solitary creatures) help to constrain most cases of ambiguity. Speakers can communicate just fine without grammatical reference to things like plurality. One could make the case that grammatical number is most relevant for human nouns, since speakers tend to talk about human referents, and since humans can occur in varying group sizes (Everett, 2019). The global distribution of grammatical number types supports this intuition: Haspelmath (2013) found that about 7% (n = 20) of the world's languages have plural marking that is optional but can only be used to denote plural human referents. Furthermore, in about 14% (n = 40) of the sampled languages, plural marking is obligatory but is restricted to human nouns. And in 5% (n = 15) of the languages, plural marking occurs on all noun types but is optional for inanimate nouns.

Haspelmath's survey reveals, then, just how variable grammatical number marking is across the world's languages. Less than half of the languages in his sample, or 46% (n = 133), exhibit the kind of grammatical number marking evident in English and most European languages, in which multiple referents must be designated with plural-marked nouns in an obligatory manner. In over half of the world's languages, grammatical plural marking is either absent, or is optional, or is only obligatory for nouns that refer to human referents. This variability of grammatical plural marking is evident across diverse regions and language families.

One logical question that follows from the diversity of grammatical number is whether one's native language impacts how s/he becomes familiar with the distinction between the notions of "one" vs. "more than one." (This topic has been raised in contemporary discussions of "linguistic relativity"; see for example Everett, 2013.) Such an impact may seem implausible given that these are quantical concepts, native to all members of our species and countless others. Yet the question is not whether variation in grammatical number enables humans' simple capacity for tracking singularity or plurality, but whether it affects how a person habitualizes themselves to such distinctions during every-day events.[1] For instance, if a person speaks a language that only indicates plurality on human nouns, does this bias that person to pay attention to quantity more when speaking about or conceptualizing human referents? Perhaps not, but to my knowledge no experimental evidence has been brought to bear on the topic. There is now evidence, however, that distinctions in grammatical number can affect how adroitly children handle quantical concepts. Some of that evidence will be discussed below.

Grammatical duals are the formal means, often noun suffixes as in the case of plural markers, that languages use to denote precisely two referents. This dual marking is not extremely rare. For instance, in a recent survey of 218 languages, Franzon et al. (2018) find that grammatical duals occur in some form in 84 of the languages. In Everett (2019) I observe that these duals tend to be restricted in terms of geographic distribution and in terms of the language families in which they occur and are also restricted in terms of function. In most languages that use dual markers, they denote distinctions on human referents only. There are over 300 language families in the world (Bickel et al., 2017), and in the vast majority of these grammatical duals are not present. Still, grammatical

1 Note that this is just one of many issues that could be raised vis-à-vis the interaction of language, culture, and cognition. For more discussion on this topic, see Everett (2017) or Saxe (2012).

dual markers are more common cross-linguistically than some might assume, given that most of the world's most widely spoken languages lack grammatical duals. One notable exception to this trend is Arabic. Intriguingly, while Spanish and English and the vast majority of European languages lack a grammatical dual maker, Proto-Indo-European did apparently have one, as did ancient Greek and Sanskrit. And there are vestiges of the grammatical dual in English, notably in the words "either" and "both."

Despite their well-known tendency to have few numbers, as in a "one-two-many" system, some languages of Australia employ grammatical dual markers. Here are some examples from Dyirbal, taken from Dixon (1972: 51):

(5) *bayi Burbula miyandanyu*
"Burbula laughed."

(6) *bayi Burbula-gara miyandanyu*
"Burbula and another person laughed."

(7) *bayi Burbula-mangan miyandanyu*
"Burbula and several other people laughed."

In (7) we see that the suffix *-mangan* serves as a plural maker, denoting that multiple people are involved in the event. But this plural is only used to denote more than two people, since if there are precisely two people the *-gara* suffix is used as in example 6. (This kind of dual marker is called an "associative dual" since it refers to a specific person and exactly one other person.) While dual markers may tend to refer to human and pronominal referents, this is certainly not the case in all languages that use them. In the Sikuani language and various others, there is a suffix or other affix that refers to precisely two things. Consider these Sikuani words: *emairibü* "a yam" vs. *emairibü-nü* "yams" vs. *emairibü-behe* "two yams." The *-behe* suffix signifies that there are precisely two yams in question (Aikhenvald, 2014).

Grammatical trials are also evident in Franzon et al.'s (2018) survey. In that survey, 20 of the 218 languages have grammatical trials. However, the grammatical trial is evident in only one world region, Oceania. It is evident in clauses like the following example sentence from Moluccan:

(8)	*duma*	*hima*	*aridu*	*na'a*
	house	that	we three	own

"We three own that house" (Laidig & Laidig, 1990: 92)

The *aridu* pronoun is a first-person trial pronoun meaning "we three." Grammatical trials are generally restricted to pronouns, even more so than grammatical duals.

Given the distribution of grammatical number types, it seems fair to say that languages generally indicate a singular/plural distinction in their grammar, either with affixes attached to the noun, or with verbal affixes or other changes made to the verb that denote "agreement" with the number of items of a relevant noun. (Verbal affixes are prefixes or suffixes, in most cases, that are attached to a verb.) This singular/plural distinction is evident throughout most of the world's languages, but a substantive minority of languages do not make the distinction grammatically. Languages that refer to grammatical duals and trials are comparably rare, and the functional utility of these other categories tends to be limited.

Does the variation that exists in the world's grammatical number types impact how speakers of languages learn basic quantitative concepts like "precisely 2" and "precisely 3"? This may seem an odd suggestion given that quantical cognition allows us to differentiate 1 from 2, and 2 from 3. Yet simply because all humans are endowed with the capacity to differentiate these quantities, we cannot assume that they come to use them in the same ways and with the same dexterity, nor that the features of a language do not impact the ease with which the concepts are handled during childhood. To the contrary, there is now evidence that grammatical number has at least a modest effect on the ease with which quantical concepts are handled, at least in some contexts. English-speaking children tend to learn the word for 1 rapidly, when compared to Japanese and Mandarin speakers (Almoammer et al., 2013; Marušič et al., 2016). This may be due, at least in part, to the presence of grammatical number in English, which Mandarin and Japanese lack. Relatedly, speakers of one dialect of Slovenian that has a grammatical dual marker tend to learn the word for 2 earlier than speakers of the other languages for which comparable data are available. These include English, Russian, Japanese, and Mandarin (Marušič et al., 2016). While such results are consistent with a grammatical effect on the ease with which even quantical concepts are labeled and manipulated linguistically, the causal role of grammar is of course debatable given the host of cultural confounds entailed in such cross-cultural research.[2] One of the ways to circumvent this challenge is to examine

AU: The reference "Marušič et al. 2016" is cited in the text but is not listed in the references list. Please either delete in-text citation or provide full reference details.

2 For example, cultures that rely heavily on trade may be more likely to refer frequently to distinctions between quantities, even small quantities (Everett, 2019). In such cases, the frequency of transactions requiring precise quantities could serve as a confounding explanation, perhaps explaining the observed differences in quantitative thought that could also correlate with linguistic differences.

groups that are relatively homogenous culturally, but differ in terms of one particular linguistic feature. Slovenian presents a critical test case, as dialects of Slovenian vary according to the presence of grammatical duals. Recent research with speakers of these dialects suggests that the kind of grammatical number that exists in a given Slovenian population impacts how and when Slovenian children learn to label and manipulate quantical concepts.

In dialects of Slovenian that employ a grammatical dual, it takes the form evident in (9).

(9)	*dva*	*rdeča*	*gumba*	*ležita*	*na*	*mizi*
	two	red.DUAL	button.DUAL	lie.DUAL	on	table

"Two red buttons are lying on the table" (Marušič et al. 2016: 2).

AU: We have changed form (10) to (9) and so on to maintain squence order. Please confirm.

Note the pervasiveness of the grammatical dual in such a clause. The adjective ("red"), the noun ("button"), and the verb ("lie") are all inflected in a way that indicates the fact that there are precisely two buttons. Learning a language like this requires children to consistently refer to whether or not there are two, and precisely two, referents being discussed. This cognitive fixation might have some effect on the age at which children become comfortable with a more general ability to symbolically denote the notion of two. A research team led by Franc Marušič at the University of Nova Gorica, Slovenia, tested the hypothesis with young children between the ages of two and four. Their sample was large, involving nearly 300 children from three Slovenian regions. Eighty-three of these children were from Slovenska Bistrica, a region of Slovenia where the dual morphology evident in clause 9 is quite normal. Seventy-one represented Central Slovenia, another region in which the grammatical dual is used. One hundred fifty-eight children represented two other regions in which speakers do not generally use the grammatical dual: Metlika and Nova Gorica. Finally, a control population of 79 English speakers in San Diego was tested. The tasks involved in the work are common to research on the development of numerical cognition. A key task was the so-called Give-N task, in which children are tested on their familiarity with basic number words. For this variant of the task, the researchers gave kids 10 buttons and asked the kids (in Slovenian or English) the following question: "Can you put *N* in the box?" For example, "Can you put two in the box?" *N* refers to a number word. The results of the Give-N task were promising for the hypothesis, pointing to subtle but significant differences across the populations of Slovenian speakers. The researchers found that "overall, speakers of dual dialects were more likely to be 2-knowers than speakers of non-dual dialects" and reached the "2-knowing" stage at an earlier age (Marušič et al., 2016: 2). While the cross-population differences were not pronounced, they were consistent with the hypothesis that grammatical dual marking can impact how kids acquire basic numbers, even those associated with quantical concepts. These and other related findings led Marušič et al. (2016) to the following conclusion: "morphological marking of number in language facilitates learning of early number word meanings" (Marušič et al., 2016: 15).

We have seen in this section that languages vary in terms of how, and whether, they denote quantitative concepts grammatically. This survey has not been comprehensive, and for a fuller picture on grammatical number I refer the reader to Corbett (2000). Yet the survey was sufficient to demonstrate that variation in grammatical number is more substantive than some scholars may presume. Furthermore, I have highlighted recent work that suggests that variation in grammatical number, including the presence/absence of a grammatical dual, may impact when and how kids are able to symbolically represent quantical concepts like 2.

3 Cross-population differences in number words

There are many critical stages in the acquisition of basic numerical concepts. These include the well-known stage at which children master the cardinal principle, becoming fully aware that a set labeled by a word *N* corresponds to an exact quantity that is associated with the word *N* only. Relatedly, they learn the successor principle, becoming aware that each word in a sequence of number words refers to the quantity denoted by the previous number word plus exactly one more (Carey, 2009a, 2009b). Prior to the acquisition of these principles, kids are able to recite a list of number words but are unaware of the relationship between them. They merely recognize that number words, like the letters of the alphabet, come in a predictable order. Much debate remains as to how exactly kids acquire the cardinal and successor principles, but it is clear that cultural variation in finger counting and number words impinge on that acquisition. The presence of precise number words like "two" or "seven" (as opposed to "few" or "several") in a language appears critical to even more basic cognitive stages that do not rely exclusively on quantical capacities. For example, the mere recognition of one-to-one correspondence benefits from the presence of number words. There is some debate as to the extent of that benefit, but work among anumeric Nicaraguan homesigners, largely anumeric Munduruku indigenes, and totally anumeric Pirahã indigenes points in the same general direction: Number words are critical to scaffolding or at least enhancing the recognition of one-to-one correspondence for set sizes larger than 3–4 (Pica et al., 2004; Spaepen et al., 2011).

For differing views on the extent of the effects of an absence of number words in a culture, see Everett and Madora (2012) and Frank et al. (2008). (The Pirahã language is well known to lack precise number words (Everett, 2005).)

Much has been written about the cultures and languages with few or no number words, and admittedly the sparse studies carried out among the relevant groups leave room for multiple interpretations of a few key results. (See Frank et al. (2008) and Everett and Madora (2012) for one example of a disagreement in interpreting the experimental results among the Pirahã.) This is not surprising given that there is still debate on the acquisition of numerical concepts in cultures whose numerical cognition has been studied with thousands of studies, for example, Americans. (See, for instance, the differing views on some key topics by prominent researchers such as Carey (2009a), and Dehaene (2011).) But it is difficult to contest that number words are critical to the acquisition of very basic numerical concepts besides the cardinal and successor principles. This conclusion is, in a way, unsurprising. What is more contestable is whether current differences in types of number words impact numerical cognition. Setting aside the rare contemporary cases of anumeric or nearly anumeric cultures, then, what can we say about the vast majority of the world's 7000+ languages that have lexical numbers? Do cultures that rely on distinct kinds of number systems exhibit associated distinctions in how they think about and learn numerical concepts? The truly cross-cultural work on this topic is modest in scope, but it does hint that variation in number word systems yields some effects on basic numerical cognition.

Anecdotally, my own impression is that the extent of diversity in the world's number systems is underestimated by many scholars. In a detailed survey of 196 languages representing dozens of families and all major geographic regions, linguist Benard Comrie offers us a sense of that diversity. Twenty of these languages have "restricted" number systems, one of which is the aforementioned extreme case of Pirahã. Other restricted cases include Hup, which will be discussed below, and some other Amazonian and Australian languages. In New Guinea there are four languages from Comrie's (2013) survey that use an "extended body part" number system. In some of these cases, for example, Kobon, counting follows a trajectory up the arm (and back down the other side of the body in some languages). So the words for 1–5 are the same as the words for the fingers on the left arm, and then 6–12 are expressible via the words for the following body parts: wrist, middle of the forearm, the elbow (or, rather, the opposite side of the elbow), the upper arm, the shoulder, the collarbone, and then, lastly, the suprasternal notch (the indentation above the sternum). Such extended body part number systems, like restricted systems, have no number bases. In 172 of the 196 languages in Comrie's (2013) survey, there are bases.

Bases of verbal numbers are the key numbers around which larger numbers are structured, usually in a multiplicative fashion. For instance, English is base-10 or decimal because number words like “forty three” are constructed around “ten”: *four x ten + three*.

According to Comrie’s (2013) survey, 125 of the 196 languages examined have decimal bases, as in English, for numbers greater than 10. A smaller but sizable segment, 20 of the 196 languages, use vigesimal or base-20 numbers for higher quantities. Hybrid bases, which rely on a combination of decimal and vigesimal bases, are found in 22 of the languages. In total, then, 167 of 196 languages in the survey use some base that is derived from an obvious anatomical source. The existence of base-10 and base-20 systems owes itself, of course, to the fact that humans have 10 fingers and 20 fingers and toes. Taking Comrie’s sample as a reasonable proxy for the world’s languages, this means that about 85% of the world’s languages likely rely on digitally based numbers, and most of the other extant number systems rely on anatomical features in some other way.

One base that is rarely attested but that has shaped much of western life, in an oblique manner at least, is the base-60 system that was once used in ancient Sumeria. This system has, over the last few millennia, worked its way into various aspects of our mathematical culture, for instance the use of 360-degree arcs evident in geometry and navigation. More fundamentally, due to its adoption by the Babylonians and Greeks, it ultimately came to shape how we define units of time. The minutes of the day are simply what one arrives at if hours are divided into 60 equal units and if we divide those units by 60 a second time we get, well, “seconds” (hours are an odd by-product of the ancient Egyptian sundials that divided the daylight into 12–10 units for when the sun was up, due to the decimal Egyptian language, plus one unit for dawn and one for dusk) (Everett, 2017). Base-60 systems are also attested in the ethnolinguistic literature, at least in the Ekari language of New Guinea:

(10) *èna ma gàati dàimita Mutò*
one and ten and Sixty
“Seventy one” (Drabbe, 1952: 30).

Interestingly, the most plausible account of the genesis of base-60 systems also points to the criticality of the fingers in the origins of numbers. An attested practice in some cultures is to count the 12 lines of the non-thumb joints of the inside of one hand with the five fingers of the other hand. (See image in Everett (2017: 80).) If each added finger is used to represent the 12 lines, then the total quantity represented by five fingers is 60 (Ifrah, 2000). So while the base-60 system we used for telling time is unrelated to the decimal system that developed in Indo-European languages, it shares with it manual origins.

There are roughly 400 Indo-European languages spoken in the world today, with English and other languages well represented across the globe as first and second languages. Proto-Indo-European, spoken somewhere in the vicinity of the Black Sea over 6,000 years ago, had a decimal system as evident by reconstructed words such as **dékmt*, "ten" and **duidkmti*, "twenty" (literally "two tens") or **trihdkomth*, "thirty" (literally "three tens"). Phonetic vestiges of such number words are still evident in descendant words, like the Portuguese word *dez* ("ten") or the word *decimal* itself, both of which bear some resemblance to **dékmt* (Everett, 2017). More critically, though, the structure of Portuguese numbers, English numbers, and numbers in other contemporary Indo-European languages still carry the structure of Proto-Indo-European numbers, whereby 10 is multiplied by smaller numbers to create larger number words. This decimal base is evident in the world's other largest language families today, including Niger-Congo, Austronesian, and Sino-Tibetan, which like Indo-European has over 400 languages and over a billion speakers. (The Niger-Congo and Austronesian families each have over 1,000 members, representing a sizable chunk of the world's 7,000+ languages.)

The manual/digital origins of number words are not simply evident in the preponderance of decimal and vigesimal number systems; they are also evident in the base-5 nature of number words less than 10 in many cultures. The critical nature of a word for 5 in constructing greater numbers is evident worldwide, and stems from the clear derivation of that number from counting with the fingers. For instance, the word for 5 in many languages is transparently derived from the word for "hand." In Proto-Austronesian, for example, the word for *hand* and the word for *five* were both *lima. The same correspondence is evident in very many unrelated languages, and the word for "five," once derived from the word for "hand," seems to kick-start the growth of larger number systems (Bowern & Zentz, 2012).

The digital foundations of numbers are even evident in some languages that have modest number systems, in words for precise numbers less than 5. In Hup and Dâw, two closely related languages of Amazonia, words for numbers are based around the kinship terms in the language. The word for 3, for instance, translates to "without a sibling" because 3 is odd. The word for 4 translates to "with a sibling," because it is even (Epps, 2006). These number words are not used by themselves, however, but alongside finger-counting strategies. So one needs to hold up four fingers and say "with a sibling" to fully denote the number four. Languages like Hup and Dâw drive home the general theme of this section: A survey of the world's spoken numbers suggests that languages vary tremendously in terms of the kinds of numbers they use, and in terms of the range of quantities denoted by those numbers. Yet there are also pervasive tendencies underlying this variability, and those tendencies point again and again to the ways in which finger counting is critical to the historical acquisition of numbers in diverse and unrelated cultural lineages.

Variation in kinds of cardinal numbers is just one of the sorts of variation in cultures' verbal representation of quantities. Ordinal numbers also vary in marked ways. In a recent survey of 321 languages, Stolz and Veselinova (2013) observe that over 10% do not have a distinct category of ordinal numbers. This is in contrast to languages like English, in which ordinal numbers are often denoted with a *-th* suffix, for example, fourth, fifth, sixth. In most languages there is some distinction between cardinal and ordinal numerals, however, and in most cases ordinal numbers are clearly derived from cardinal numbers as in the English examples just cited. Intriguingly, though, in almost two thirds of the languages surveyed by Stolz and Veselinova (2013), small ordinal numbers are treated differently. In many of these languages it is only the ordinal number for 1, as in English "first" (we do not say "oneth')" In some languages 2 also is denoted with a distinct ordinal number, as with English "second" (we do not say "twoth"). The cross-cultural variation in small ordinal numbers underscores that even basic reference to quantical concepts (quantities less than four) varies cross-culturally. This variation in the reference to quantical concepts, which was also evident in our discussion of cardinal numbers and grammatical numbers, is in some sense surprising. Languages vary extensively with respect to how they describe quantical concepts that all humans share and, as seen in cases like Slovenian, this variation has demonstrable effects on the age at which individuals become adept and using such "quantical" concepts. While linguists, anthropologists, psychologists, and others have long been aware of variation in terms of how languages denote numerical concepts, only relatively recently have we come to appreciate that that variation extends in key ways to quantical concepts. It is possible, however, that we still underestimate the ways in which languages vary vis-à-vis their expression of quantical concepts. In a very recent study involving data from nearly 6,000 dialects, I make the case that there is another key type of variation in number words for quantical concepts that has still not been explored systematically: The cross-cultural frequency in speech of words for 1 and 2 (Everett, 2019).

While the vast majority of the world's languages have words translatable as "one" and "two," this does not mean that those terms are used in the same ways or at the same rate. The exploration of their frequency seemed worthwhile for a few reasons. One reason is that the frequency of usage of number terms, even as small as "one" and "two," could well impact the rate and age at which children become practiced with basic quantitative concepts. This possibility is supported by the aforementioned work on grammatical duals, which suggests that the frequent grammatical reference to 2 facilitates to some degree children's refinement of certain facets of basic quantitative thought. While directly establishing the frequency in speech of words like "two" for most of the world's languages is not possible, there is one indirect way to test for frequency in speech. This way relies on a well-known fact about words: Highly frequent words tend to be reduced phonetically, that is, made shorter (Bybee, 2007). With this fact in mind, I examined the length of number words for "one" and "two" across the bulk of the world's languages. This was done via a database containing 40–100 commonly used words (phonetically transcribed) for the bulk of the world's languages (Wichmann et al., 2018). My work looked at 5,942 language varieties (dialects and mutually unintelligible languages), considering the average word length of all the words for each language. For each language variety, I then contrasted the word lengths for "one" and "two," respectively, with the average word length of all the other words in that language. Upon doing so, a very clear pattern emerged: The languages spoken by cultures with larger populations tend to have shorter words for "one" and "two," even after controlling for factors like the average word lengths of particular languages and the relatedness of languages. This pattern suggests strongly that larger populations tend to use number words more frequently than smaller populations. There are many factors that likely motivate this tendency across the world's culture, including greater frequency of number words in cultures relying on trade and industrialization.

This all may seem very intuitive and even trivial: Of course cultures vary in the degree to which they use number words, and in the frequency with which they use number words in practices like trade. Yet the key point is that such variation extends to number words for quantical concepts that are shared by all human populations. Previous work had suggested that quantical concepts, namely 1, 2, and 3, are less prone to being concretized in varied ways across cultures because they are native concepts (Franzon et al., 2018). Instead, I argue, they are treated pretty much like other quantitative concepts in terms of how they are referred to in speech. That is, they are prone to cross-cultural variation and are used with very different frequencies across the world's cultures – at least judging from the indirect word-length data. More broadly, the issue of the frequency of small number words raises yet another kind of cross-linguistic variation in numbers. This variation, like the variation in grammatical number types, may impact children's acquisition of numerical concepts. Work is required to explore this possibility.

In this section we have seen that there is an underlying manual basis of number systems but also an amazing diversity of number words overlaid over that manual basis. This includes diversity of several sorts: Diversity in number bases (despite their generally digital origins), diversity in the mere existence of number words (since some languages lack them), diversity in ordinal numbers, and diversity in the frequency with which numbers, even very small numbers, are used. This global diversity of number words impacts how kids acquire numerical concepts and even their facility with basic quantical concepts. All these factors are worth keeping in mind when considering how best to teach arithmetic across the world's cultures. The linguistic features of a given culture affect how the members of that culture learn even basic quantitative concepts.

4 Discussion and conclusion

While there are universal human quantitative capacities, each culture and language brings with it its own biases in terms of how it refers to quantical and numerical concepts. A greater awareness of the extant cross-cultural diversity of spoken numbers could, I hope, benefit those concerned with how best to teach basic arithmetic concepts. It is still very debatable just how much cross-cultural variation of numbers impacts how kids acquire numerical concepts. Yet, where relevant experimental evidence exists, it consistently suggests that such variation matters, often in marked ways. If people speak an anumeric language, this has marked effects on their ability to learn number concepts. If they speak a language with a grammatical dual, this seems to offer advantages to early numerical cognition. More commonly, cross-linguistic variation in the transparency of number bases may impact how kids acquire numbers. Some evidence suggests that Chinese children, for instance, outperform children from the UK, Russia, and other nations on mathematical tasks, and that this high performance is due in part to the greater transparency of the decimality of Chinese numbers (Rodic et al., 2015; though see Moschkovich, 2017). So, while languages tend to have decimal bases, the transparency with which decimality is expressed appears to affect the cross-cultural acquisition of numerical concepts.

All of this leaves us with two simple conclusions: (1) The cross-cultural variation of linguistic numbers impacts quantitative cognition, and (2) the cross-cultural variation of linguistic numbers is remarkable even if it is constrained by the typically digital origins of numbers. Both of these points seem worth bearing in mind as we adopt and refine pedagogical models for arithmetic instruction, if we are interested in the cross-cultural efficacy of those models.

References

Aikhenvald, Alexandra (2014): Number and noun categorisation: A view from north-west amazonia. In Storch, Anne, Dimmendaal, Gerrit J. eds.: *Number – Constructions and Semantics. Case Studies from Africa, Amazonia, India and Oceania*. Amsterdam: John Benjamins. ~~xxx–xxx.~~

Almoammer, Alhanouf et al. (2013): Grammatical morphology as a source of early number word meanings. *Proceedings of the National Academy of Sciences of the United States of America* 110, 18448–18453.

Bickel, Balthasar, Nichols, Johanna, Zakharko, Taras, Witzlack-Makarevich, Alena, Hildebrandt, Kristine, Rießler, Michael, Bierkandt, Lennart, Zúñiga, Fernando & Lowe, John B. (2017). *The AUTOTYP typological databases*. Version 0.1.0. doi: DOI 10.5281/zenodo.3667562.

Bowern, Claire & Zentz, Jason (2012): Diversity in the numeral systems of Australian languages. *Anthropological Linguistics* 54, 133–160.

Bybee, Joan (2007): *Frequency of Use and the Organization of Language*. Oxford: Oxford University Press.

Carey, Susan (2009a): *The Origin of Concepts*. Oxford: Oxford University Press.

Carey, Susan (2009b): Where our number concepts come from. *The Journal of Philosophy* 106, 220–254.

Comrie, Bernard (2013): Numeral bases. In Dryer, Matthew S., Haspelmath, Martin eds.: *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. ~~xxx–xxx.~~

Corbett, Greville (2000): *Number*. Cambridge: Cambridge University Press.

Dehaene, Stanislas (2011): *The Number Sense*. Oxford: Oxford University Press.

Dixon, Robert & Ward, Malcolm (1972): *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.

Drabbe, Peter (1952): *Spraakkunst Van Het Ekagi*. 'S-Gravenhage: Martinus Nijhoff.

Epps, Patience (2006): Growing a numeral system. *Diachronica* 23, 259–288.

Everett, Caleb (2013): *Linguistic Relativity: Evidence Across Languages and Cognitive Domains*. Berlin: De Gruyter Mouton.

Everett, Caleb (2017): *Numbers and the making of us: Counting and the Course of Human Cultures*. Cambridge, MA: Harvard University Press.

Everett, Caleb (2019): Is native quantitative thought concretized in linguistically privileged ways? A look at the global picture. *Cognitive Neuropsychology* 10.1080/02643294.2019.1668368.

Everett, Caleb & Madora, Keren (2012): Quantity recognition among speakers of an anumeric language. *Cognitive Science* 36, 130–141.

Everett, Daniel (2005): Cultural constraints on grammar and cognition in Pirahã. *Current Anthropology* 46, 621–646.

Frank, Michael, Everett, Daniel L., Fedorenko, Evelina & Gibson, Edward (2008): Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition* 108, 819–824.

Franzon, Francesca, Zanini, Chiara & Rugani, Rosa (2018): Do non-verbal systems shape grammar? Numerical cognition and number morphology compared. *Mind & Language* 10.1111/mila.12183.

AU: Please update page number in Aikhenvald (2014).

AU: Please update page number in Comrie (2013).

ease update page number in Haspelmath (2013).

Haspelmath, Martin (2013): Coding of nominal plurality. In Dryer, Matthew S., Haspelmath, Martin eds.: *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. ~~xxx–xxx.~~

Ifrah, Georges (2000): *The Universal History of Numbers: From Prehistory to the Invention of the Computer*. Hoboken, New Jersey: Wiley.

Laidig, Wyn & Laidig, Carol (1990): Larike pronouns: Duals and trials in a central moluccan language. *Oceanic Linguistics* 29, 87–109.

Moschkovich, Judit N. (2017): Revisiting early research on early language and number names. *Eurasia Journal of Mathematics, Science & Technology Education* 13 (7b), 4143–4156.

Núñez, Rafael (2017): Is there really an evolved capacity for number. *Trends in Cognitive Sciences* 2, 409–424.

Pica, Pierre, Lemer, Cathy, Izard, Veronique & Dehaene, Stanislas (2004): Exact and approximate arithmetic in an Amazonian indigene group. *Science* 306, 499–503.

Rodic, Maya, Zhou, Xinlin, Tikhomirova, Tatiana, Wei, Wei, Malykh, Sergei, Ismatulina, Victoria, Sabirova, Elena, Davidova, Yulia, Tosto, Maria Grazia, Lemelin, Jean-Pascal & Kovas, Yulia (2015): Cross-cultural investigation into cognitive underpinnings of individual differences in early arithmetic. *Developmental Science* 18 (1), 165–174.

Saxe, Geoffrey (2012): *Cultural Development of Mathematical Ideas: Papua New Guinea studies*. Cambridge, UK: Cambridge University Press https://doi.org/10.1017/CBO9781139045360.

Spaepen, Elizabet, Coppola, Marie, Spelke, Elizabeth S., Carey, Susan E. & Goldin-Meadow, Susan (2011): Number without a language model. *Proceedings of the National Academy of Sciences* 108, 3163–3168.

AU: Please update page number in Stolz, & Veselinova (2013).

Stolz, Thomas & Veselinova, Ljuba (2013): Ordinal numerals. In Dryer, Matthew, Haspelmath, Martin eds.: *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. ~~xxx–xxx.~~

Wichmann, Soren, Holman, Eric & Brown, Cecil (eds.) (2018). The ASJP database (version 18).