Numeral systems across languages support efficient communication: From approximate numerosity to recursion

Yang Xu (yang_xu_ch@berkeley.edu)
Terry Regier (terry.regier@berkeley.edu)
Language and Cognition Lab, Department of Linguistics
University of California, Berkeley, CA 94720 USA

Abstract

Languages differ qualitatively in their numeral systems. At one extreme, some languages have a small set of number terms, which denote approximate or inexact numerosities; at the other extreme, many languages have forms for exact numerosities over a very large range, through a recursively defined counting system. What explains this variation? Here, we use computational analyses to explore the numeral systems of 25 languages that span this spectrum. We find that these numeral systems all reflect a functional need for efficient communication, mirroring existing arguments in the domains of color, kinship, and space. Our findings suggest that cross-language variation in numeral systems is shaped by the functional need to communicate precisely, using minimal cognitive resources.

Keywords: number; semantic variation; efficient communication; functionalism; recursion; language and thought.

Numeral systems and recursion

A central question in cognitive science is why languages partition human experience into categories in the ways they do (Berlin & Kay, 1969; Levinson & Meira, 2003). Another central question concerns the basis and nature of recursion (Hauser, Chomsky, & Fitch, 2002; Pinker & Jackendoff, 2005). These two questions converge in the study of *numeral systems* across languages.

Languages vary in their numeral systems (Greenberg, 1978; Comrie, 2013). Moreover, there are qualitatively distinct classes of such numeral systems. Some languages have numeral systems that express only approximate or inexact numerosity; other languages have systems that express exact numerosity but only over a restricted range of relatively small numbers; while yet other languages have fully recursive counting systems that express exact numerosity over a very large range. This increasing precision from approximate to exact systems is reflected in child development: at age 3 or 4, children who are learning an exact counting system transition from an approximate to a precise understanding of number words (Wynn, 1990). This "crystallization" of discrete numbers out of an approximate base has been argued to be just what adult speakers of languages with approximate numeral systems do not have, compared with those of exact ones (Pica et al., 2004; Gordon, 2004). Instead, approximate numeral systems appear to be grounded directly in the non-linguistic approximate number system, a cognitive capacity for approximate numerosity that humans share with non-human animals (Dehaene, 2011).

We seek to understand why certain numeral systems are attested in the world's languages while other logically possible systems are not. We also seek to understand why the qualitative classes of such systems – from approximate numerosity, to exact counting over a restricted range, to fully recursive counting – appear as they do.

Efficient communication

An existing proposal has the potential to answer these questions. It has been argued that systems of word meanings across languages reflect the need for efficient communication. On this account, for any given semantic domain, the different categorical partitionings of that domain observed in the world's languages represent different means to the same functional end: communicating precisely, with minimal expenditure of cognitive resources. This idea is supported by cross-language computational analyses in the domains of color (Regier et al., 2007), kinship (Kemp & Regier, 2012), and space (Khetarpal et al., 2013) – and it also reflects a more general recent focus on efficient communication as a force that explains why languages take the forms they do (Fedzechkina et al., 2012; Piantadosi et al., 2011; Smith et al., 2013). We ask here whether the same idea explains why numeral systems appear as they do, from approximate to fully recursive form.

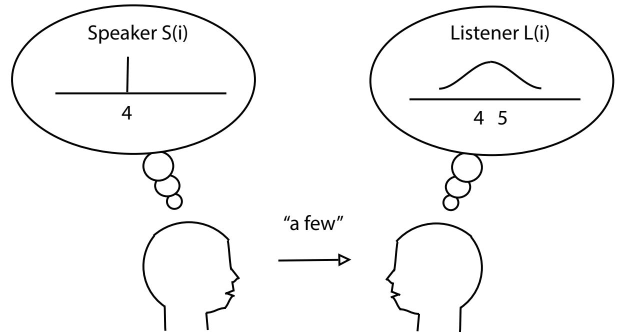

Figure 1: Scenario for communicating a number.

The idea of efficient communication involves a tradeoff between two competing forces: informativeness and simplicity. An informative system is one that supports precise communication; a simple system is one with a compact cognitive representation. A maximally informative system would have a separate word for each object in a given semantic domain – which would be complex, not simple. In contrast, a maximally simple system would have just one word for all objects in a given semantic domain – which would not support precise communication. The proposal is that attested semantic systems are those that achieve a near-optimal tradeoff between these two competing principles, and thus achieve communicative efficiency (Kemp & Regier, 2012).

Figure 1 illustrates these ideas. [1] Here, a speaker has a particular number in mind (4, mentally represented as an exact point on a number line), and wishes to convey that number to a listener. The speaker has expressed that number using the English approximate term "a few", rather than the exact term "four" that is also available in English. On the basis of this utterance, the listener *mentally reconstructs* the number that the listener believes the speaker intended. Because the term "a few" is inexact, the listener's reconstruction of the intended number is also inexact, and is shown as a probability distribution centered near 4 or 5 and extending to neighboring numbers as well. As a result, the listener's mental reconstruction does not match the speaker's intention perfectly. However, if the speaker had instead used the exact term "four", that term would have allowed the listener to reconstruct the speaker's intended meaning perfectly. We take the informativeness of communication to be the extent to which the listener's mental reconstruction matches the speaker's representation. Communication is not perfectly informative in the case of "a few" but would be perfectly informative in the case of "four".

Clearly, an exact numeral system that picks out specific integers is more informative than an approximate system - but it is less simple. A system of approximate numerals can span a given range of the number line using very few terms, whereas many exact integer terms would be needed to span the same range. Thus the high informativeness of an exact numeral system comes at a high cognitive cost. Importantly, however, a *recursive* exact system would be specifiable using a relatively small number of generative rules, rather than separate lexical entries for each exact numeral. Thus, recursive numeral systems may be a cognitive tool (Frank et al., 2008) that enables highly informative communication about number at the price of only modest cognitive complexity.

We wish to know whether these ideas can account for which numeral systems, and which classes of such systems, are attested across languages. To test this, we require: (1) a dataset of cross-language numeral systems, (2) a formal specification of our theory, and (3) a test of the theory against the data. We specify each of these in turn below, and then present our results. To preview those results, we find that numeral systems across languages strongly tend to support near-optimally efficient communication, and that the drive for efficient communication also helps to explain why the different classes of numeral systems appear as they do. Our results suggest that the different types of numeral system found across languages all support the same functional goal of efficient communication, in different ways.

Data

We considered the numeral systems of 25 languages. All but one were from Comrie (2013), a survey of numeral bases in the *World Atlas of Language Structures* (WALS). This survey includes references to grammars for individual languages, each of which describes that language's numeral system. Comrie (2013) draws a distinction between "restricted" numeral systems, which he defines as those that do "not effectively go above around 20", and other numeral systems, which cover a larger range, often through recursion. We focused on all 20 languages that Comrie had listed as "restricted", together with several representative languages from the same source that have numeral systems over a larger range. These numeral systems from WALS were supplemented by a description of the Mundurukú numeral system from Pica et al. (2004); the data from this latter source are more detailed than those from the grammars.

These languages span the spectrum from approximate, to exact restricted, to fully recursive numeral systems. [2] We have used these class designations loosely up until now; we now define them precisely. We took a language's numeral system to be *approximate* if the grammar or other description on which we relied for that language explicitly states that the meanings of the numerals in the system are approximate or inexact. All such systems in our data were restricted in Comrie's sense. We took a language's numeral system to be *exact restricted* if the system covers a restricted range (again in Comrie's sense) but the description of the system does not explicitly state that the meanings were approximate or inexact; thus we assumed exactness unless there was evidence to the contrary. Finally, we took a language's numeral system to be *recursive* if the numeral system was listed by Comrie as having a base: e.g. the English numeral system is recursive with base 10. [3] These categories do not entirely exhaust the space of attested systems. For example, Comrie lists several extended body-part numeral systems, which use body parts beyond the 10 fingers to enumerate, and can reach well above 20, and there are some restricted languages that use recursion within a limited range. However, the three broad classes we use do pick out major types of numeral system.

Formal presentation of theory

We have seen that the notion of efficient communication involves a tradeoff between the competing forces of simplicity and informativeness. We first formalize each of these two forces in turn, and then the tradeoff between them. This formalization builds on that of Kemp & Regier (2012).

Simplicity

Simplicity is the opposite of complexity, and we define the *complexity* of a numeral system to be the number of rules

[1] The symbols S(i) and L(i) in the figure are explained in our formal presentation below.

[2] Approximate: Pirahã, Wari (3 terms); Gooniyandi (5 terms); Mundurukú (6 terms). Exact restricted: !Xóõ, Achagua, Araona, Baré, Hixkaryána, Mangarrayi, Martuthunira, Pitjantjatjara (4 terms); Awa Pit, Kayardild (5 terms); Barasano, Hup, Imonda, Rama, Yidiny (6 terms); Waskia, Wichí (11 terms). Recursive: Chinese, English (base-10); Ainu (base-20); Georgian (base-10 & 20).

[3] We focus on the most fine-grained set of numeral terms available in each language, ignoring for now approximate terms in languages with an exact numeral system, e.g. "a few" in English.

needed to specify it (Kemp & Regier, 2012). We specify numeral systems as grammars, based on the primitive components listed in Table 1.

Table 1: Components for representing numeral systems.

Component	Description
1	Primitive concept of 1
m(w)	Meaning of form w
g(x)	Gaussian with mean x
s(w,n)	Successor of w with length n
S = {}	Specification of set S
+	Addition
-	Subtraction
×	Multiplication

Two of these primitives require explanation. g(x) is a Gaussian centered at position x on a number line that scales in accord with the non-linguistic approximate number system; it is intended to ground approximate numerals directly in that cognitive system. s(w,n) is a generalization of the standard successor function (successor(w)= m(w)+1); it defines a line segment that begins at m(w)+1 and continues for some exactly specified length n, picking out the interval [m(w)+1,m(w)+n]. Although in attested systems this length n is generally 1, the more general case is used for hypothetical numeral systems against which we compare attested ones.

Table 2 presents grammars for the numeral systems of three languages, one from each of the three classes we consider here, and indicates the complexity of each. Here and elsewhere in this paper, we restrict our attention to numerals over the range 1-100.

Informativeness

Informativeness of communication was illustrated in the communicative scenario of Figure 1. Returning to that scenario, we may represent the speaker's intended meaning as a probability distribution $S(i)$ over numbers i, and analogously represent the listener's mental reconstruction of that meaning as a distribution $L(i)$ over numbers i. We assume that the speaker has certainty: $S(t) = 1$ for the intended target number t. We assume moreover that the character of the listener distribution $L(i)$ depends on the number term produced by the speaker. If the speaker has produced a word w that is semantically grounded in the approximate number system (Dehaene, 2011) via a rule involving the primitive g($\cdot$), we assume that the listener distribution takes the form:

$$L(i) \propto f(i|w) = \exp\left[-\left(\frac{|i-\mu_w|}{\sqrt{2}v\sqrt{i^2+\mu_w^2}}\right)^2\right] \qquad (1)$$

This equation is derived from the first equation on p. S6 of the supporting online materials for Pica et al. (2004), who present it as a formalization of the comparison of two numerosities, both represented as Gaussians, in the non-linguistic approximate number system, which obeys Weber's law. $f(i|w)$ captures the similarity of the number i to the mean μ_w of the category named by the word w, under the assumptions of Weber's law. v is the empirically determined Weber fraction, which we take to be 0.17 in our analyses, following Pica et al. (2004). [4]

Table 2: Grammars for sample *approximate*, *exact restricted*, and *recursive* numeral systems over the range 1-100. The specification of each rule adds a unit complexity of 1.

	Approximate: Pirahã	Complexity: 3
Number	*Form*	*Rule*
1	'hói'	g(1)
∼2-4	'hoí'	g(2)
∼5-100	aibaagi	g(7)
	Exact: Kayardild	Complexity: 7
Number	*Form*	*Rule*
1	'warngiida'	1
2	'kiyarrngka'	s^1('warngiida')
3	'burldamurra'	s^1('kiyarrngka')
4	'mirndinda'	s^1('burldamurra')
5-100	'muthaa'	s^∞('mirndinda')
		$s^1(w)$ = s(w,1)
		$s^\infty(w)$ = s(w,∞)
	Recursive: English	Complexity: 24
Number	*Form*	*Rule*
1	'one'	1
2	'two'	s^1('one')
3-12	'three'...'twelve'	s^1('two')...s^1('eleven')
20	'twenty'	m('two')×m('ten')
100	'hundred'	m('ten')×m('ten')
13...19	w'teen' : $w \in$ $\text{ones}_{\setminus 1,2}$	m(w) + m('ten')
30...90	w'ty' : $w \in$ $\text{ones}_{\setminus 1,2}$	m(w)×m('ten')
{21...29 91...99}	u-v : $u \in$ tens, $v \in$ ones	m(u) + m(v)
		$s^1(w)$ = s(w,1)
		$\text{ones}_{\setminus 1,2}$ = {'thir','for','four', 'fif','six'...'nine'}
		'thir' = 'three'
		'for' = 'four'
		'fif' = 'five'
		ones = {'one'...'nine'}
		tens = {'twenty'...'ninety'}

In contrast, if the speaker has used an exact number term w grounded in exact primitives such as s($\cdot,\cdot$), we assume that the listener distribution is given by the size principle (Tenenbaum, 1999):

$$L(i) = p(i|w) = \frac{1}{|w|} \qquad (2)$$

[4] We also repeated our analyses with v = 0.3 at the suggestion of Stanislas Dehaene (personal communication); the results did not differ qualitatively from those we report below.

where $|w|$ is the number of integers contained in the exact interval named by the number word w. [5] In the case of most attested systems, an exact numeral such as "nine" will pick out just a single integer, so that $p(9|\text{"nine"}) = \frac{1}{1} = 1$. However the formula also generalizes to hypothetical exact numerals defined as longer exact intervals of the number line.

Given these specifications of the speaker (S) and listener (L) distributions, we define the communicative cost $C(i)$ of communicating a number i under a given numeral system to be the *information lost* in communication – that is, the information lost in the listener's reconstruction L when compared to the speaker's distribution S. We model this information loss as the Kullback-Leibler (KL) divergence between distributions S and L. In the case of speaker certainty ($S(i) = 1$ for the target number i), this reduces to surprisal:

$$C(i) = D_{KL}(S||L) = \sum_i S(i) \log_2 \frac{S(i)}{L(i)} = \log_2 \frac{1}{L(i)} \quad (3)$$

We model the communicative cost for a numeral system as a whole as the expected value of C over all numbers i:

$$E[C] = \sum_i N(i)C(i), \quad (4)$$

Here, $N(i)$ is the need probability of target number i. We estimated need probabilities by the normalized frequencies of English numerals in the Google ngram corpus (Michel et al., 2011) for the year 2000, smoothed with an exponential curve via log-linear regression (Pearson correlation with unsmoothed data $= 0.97$). [6]

Tradeoff

We take a numeral system to be simple to the extent that it exhibits low complexity, and we take it to be informative to the extent that it exhibits low communicative cost $E[C]$. Given this, we consider a numeral system to be *near-optimally efficient* if it is more informative (i.e. exhibits lower communicative cost) than most logically possible hypothetical systems of the same complexity.

Testing the theory

We test our theory against the data in two steps. We first test whether the model of the approximate number system represented by Equation 1 accommodates fine-grained linguistic data from the one language for which we have such data, Mundurukú. We then test whether all numeral systems in our dataset are near-optimally efficient in the sense defined above.

Mundurukú and the approximate number system

Pica et al. (2004) showed that their formalization of the approximate number system, governed by Weber's law, accounted well for *non-linguistic* numerosity judgments by the Mundurukú. They also collected fine-grained data on the way the Mundurukú *name* different numerosities, but they did not directly test whether their formalization of the approximate number system also accommodates those linguistic data.

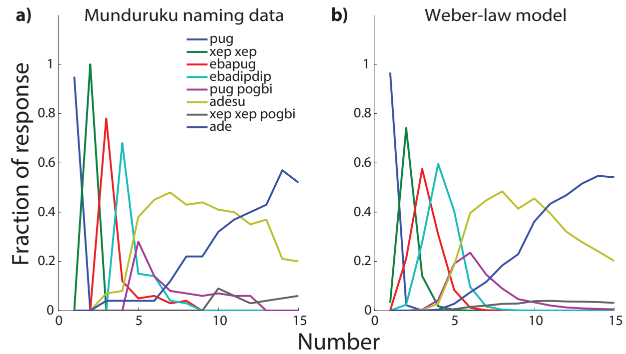

Figure 2: Modeling Mundurukú naming data.

Figure 2(a) shows, for numerosities 1 to 15, the fraction of times each numerosity i was named with a given Mundurukú word or locution w (Pica et al., 2004).[7] We modeled this fraction $p(w|i)$ using Bayes' rule: $p(w|i) \propto f(i|w)p(w)$, where $f(i|w)$ is given by Equation 1, based on Weber's law, and the prior $p(w)$ is given by the relative frequency of word w in the data, over all numerosities. We fit this model to the data in Figure 2(a) by finding placements of category means μ_w that minimize the mean-squared-error (MSE) between model and data. The model fit was good (MSE = 0.004), and is shown in Figure 2(b). A variant of this model based instead on exact numeral representation (Equation 2) performed much more poorly (MSE = 0.03, fit not shown). These findings suggest that the model of the approximate number system given by Equation 1 provides a reasonable basis for grounding approximate numeral systems.

Near-optimal efficiency of numeral systems

To test whether the attested numeral systems in our dataset are near-optimally efficient, we assessed their simplicity and informativeness relative to a large set of logically possible hypothetical systems. These hypothetical systems fell in the same three classes as our attested systems: approximate, exact restricted, and recursive.

We generated approximate hypothetical systems that have $k = 3$ through $k = 20$ numeral categories, and that place these categories at the lower end of the number line, specifically in the interval [1,20]. We did this because these specifications also accommodate the attested systems in our dataset, and we wanted to generate hypothetical systems that were broadly comparable to actual ones. For each k, we exhaustively enumerated all possible placements of k means for a k-term system in the interval [1,20], producing $\binom{20}{k}$ systems

[5] $L(i) = 0$ if i lies outside the category named by w.

[6] $N(i)$ decays roughly exponentially with increasing i. Data from spoken English (Leech et al., 2001) show a similar trend.

[7] The data for a given numerosity i do not always sum to 1.0 because some infrequent terms were not reported. In our model, we accommodated this fact by introducing an unnamed dummy category corresponding to these unreported terms.

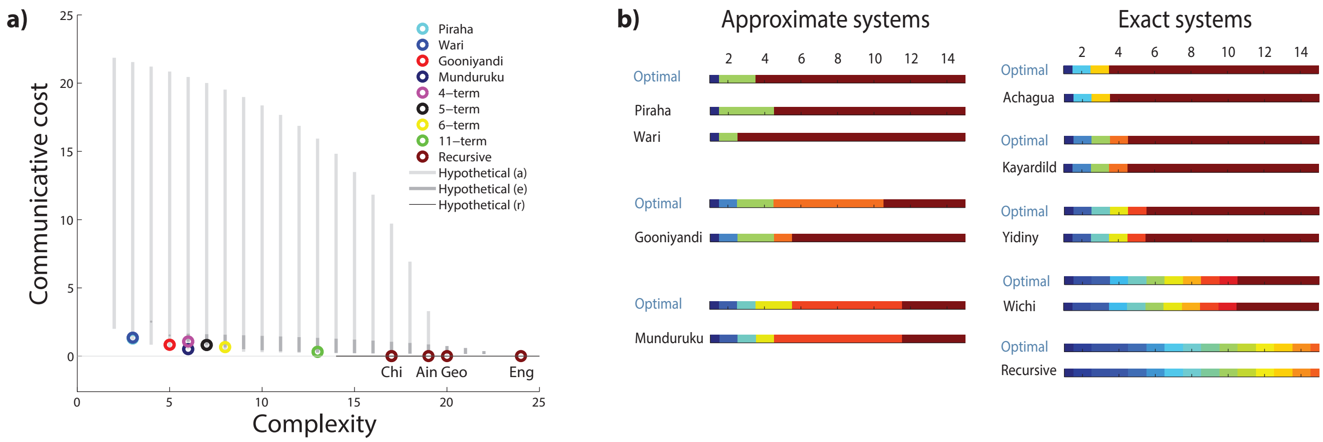

Figure 3: (a) Near-optimal tradeoff between communicative cost and complexity across 25 attested numeral systems, compared with hypothetical approximate (a), exact restricted (e) and recursive (r) systems. 'Chi', 'Ain', 'Geo', 'Eng' stand for 'Chinese', 'Ainu', 'Georgian' and 'English'. (b) Comparison of sample attested systems to optimal systems of the same complexity.

for each k. The grammar for each such system was simply a list of Gaussians g(·) centered at these means, by analogy with the Pirahã grammar in Table 2.

We generated exact restricted hypothetical systems analogously. For $k = 3$ through $k = 20$, we exhaustively enumerated all possible partitions of the range [1,20] into k non-overlapping exact intervals, and took each such partition to be a hypothetical system. The grammar for each such system represented these intervals using the s(·,·) primitive. As in the Kayardild grammar in Table 2, we required an additional separate rule to define each interval length used in the grammar. As a result, exact numeral system grammars that use intervals of many different lengths are penalized in complexity. Systems with numerals that pick out specific integers are included in this class: in this case the interval size is 1.

Finally, we generated hypothetical recursive systems by considering the full space of canonical base-n recursive numeral systems (Hurford, 1999) for n = 2 to 30. We took a canonical base-n system to be one in which there are distinct lexical items for the numerals 1 through n, and all numerals beyond that are constructed by generative rules according to the pattern $xn + y$ for some already-defined numerals x, y (Comrie, 2013). In these systems, all numerals correspond to specific integers. The English grammar in Table 2 is not perfectly canonical because the teens are part of a separate subsystem from other high numerosities, but it is an example of a recursive base-10 system.

Figure 3(a) shows all hypothetical and attested systems plotted according to their complexity and communicative cost. Colored circles denote attested systems. Light gray bars denote the range of costs exhibited by approximate hypothetical systems of a given complexity; dark gray bars denote the range of costs exhibited by exact restricted hypothetical systems of a given complexity; and the extent of the black horizontal line (not including the light gray portion of that line) at communicative cost 0 denotes the range of complexities exhibited by hypothetical recursive systems, all of which have communicative cost 0. It can be seen that in general, attested numeral systems in our dataset tend to be more informative (show lower communicative cost) than most hypothetical alternatives of the same complexity. Thus, these attested systems do support near-optimally efficient communication about number. Figure 3(b) shows sample systems from our dataset compared with the theoretically optimally informative (lowest cost) systems of the same complexity – in all cases color-coded such that a numeral corresponds to a colored region of the number line. It can be seen that the attested systems resemble these theoretical optima.

These results support a functional account of why the different classes of numeral system in the world's languages appear as they do, namely as different ways of navigating the tradeoff between simplicity and informativeness. Approximate numeral systems (e.g. Wari, with 3 terms) represent one extreme on a continuum: they are simple (non-complex), requiring only a minimal cognitive investment in communicating about number. These systems support near-optimally informative communication for that level of cognitive investment – but they do not closely approach perfectly informative (0 cost) communication. Exact restricted systems (e.g. Kayardild, with 5 terms) are slightly more complex – and they support somewhat more informative communication. Finally, recursive systems represent the informative extreme of this continuum: these systems support perfectly informative communication, because there is a (recursively generated) separate name for each integer within a large range. Such fine-grained naming would be prohibitively expensive under an exact non-recursive or an approximate system: one lexical rule per integer in the range. But a recursive system can be seen as a cognitive tool that supports perfectly informative communication over a large range, at the cost of only modest complexity. Interestingly, our exploration of hypothetical recursive systems revealed (although not shown in the figure) that base-10 and base-20 systems are near-optimally simple, whereas base-2 systems are quite complex – which could explain the cross-linguistic prominence of base-10 and base-20 systems, and the very low frequency of base-2 recursive systems over a large range.

Conclusions

We have seen that the need for efficient communication helps to explain why numeral systems across languages take the forms they do, by analogy with recent demonstrations in other semantic domains – and that the same functional need helps to explain the qualitatively different classes of numeral system found across languages. At the core of this explanation is the idea that attested numeral systems near-optimally trade off the competing demands of informativeness and simplicity.

Several questions are left open by these findings. Will the results generalize to other languages? Are these findings themselves dependent on simplifying assumptions we have made? What sort of evolutionary process produces these patterns? Future studies should address these questions, to place our present findings in their proper context. For now, however, we hope that the evidence we have presented helps to explain the diversity of numeral systems in the world's languages, in terms of the functional drive for efficient communication.

Acknowledgments

We thank Charles Kemp for his role in developing the computational framework we use. We thank Stanislas Dehaene and three anonymous reviewers for their comments. This work was supported by NSF grant SBE-1041707, the Spatial Intelligence and Learning Center (SILC).

References

Berlin, B., & Kay, P. (1969). *Basic color terms: Their universality and evolution*. Berkeley: University of California Press.

Comrie, B. (2013). Numeral bases. In M. S. Dryer & M. Haspelmath (Eds.), *The world atlas of language structures online.* Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at http://wals.info/chapter/131, Accessed on 2014-01-26)

Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. New York: Oxford.

Fedzechkina, M., Jaeger, T. F., & Newport, E. L. (2012). Language learners restructure their input to facilitate efficient communication. *PNAS*, *109*, 17897–17902.

Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Pirahã language and cognition. *Cognition*, *108*, 819–824.

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

Greenberg, J. H. (1978). Generalizations about numeral systems. In J. H. Greenberg (Ed.), *Universals of human language, volume 3: Word structure* (pp. 249–295). Stanford: Stanford University Press.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, *298*, 1569–1579.

Hurford, J. (1999). Artificially growing a numeral system. In J. Gvozdanovic (Ed.), *Numeral types and changes worldwide.* Berlin: De Gruyter.

Kemp, C., & Regier, T. (2012). Kinship categories across languages reflect general communicative principles. *Science*, *336*, 1049–1054.

Khetarpal, N., Neveu, G., Majid, A., Michael, L., & Regier, T. (2013). Spatial terms across languages support near-optimal communication: Evidence from Peruvian Amazonia, and computational analyses. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual meeting of the Cognitive Science Society.*

Leech, G., Rayson, P., & Wilson, A. (2001). *Word frequencies in written and spoken English based on the British National Corpus.* London: Longman.

Levinson, S. C., & Meira, S. (2003). 'Natural concepts' in the spatial topologial domain–adpositional meanings in crosslinguistic perspective: An exercise in semantic typology. *Language*, *79*, 485–516.

Michel, J. B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., The Google Books Team, et al. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, *331*, 176–182.

Piantadosi, S. T., Tily, H., & Gibson, E. (2011). Word lengths are optimized for efficient communication. *PNAS*, *108*, 3526–3529.

Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, *306*, 499–503.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition*, *95*, 201–236.

Regier, T., Kay, P., & Khetarpal, N. (2007). Color naming reflects optimal partitions of color space. *PNAS*, *104*, 1436–1441.

Smith, K., Tamariz, M., & Kirby, S. (2013). Linguistic structure is an evolutionary trade-off between simplicity and expressivity. In M. Knauff, M. Pauen, N. Sebanz, & I. Wachsmuth (Eds.), *Proceedings of the 35th annual meeting of the Cognitive Science Society.*

Tenenbaum, J. B. (1999). *A Bayesian framework for concept learning*. Doctoral dissertation, MIT.

Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*, 155–193.