

The Number Sense and the Language Instinct: Relating Syntax and Arithmetic in Human Cognition

José-Luis Mendívil-Giró¹ 

[1] Department of Linguistics and Hispanic Literatures, University of Zaragoza, Zaragoza, Spain.

biolinguistics, 2025, Vol. 19, Article e17287, <https://doi.org/10.5964/bioling.17287>

Received: 2025-03-10 • Accepted: 2025-04-06 • Published (VoR): 2025-05-19

Handling Editor: Kleantes K. Grohmann, University of Cyprus, Nicosia, Cyprus

Corresponding Author: José-Luis Mendívil-Giró, Pedro Cerbuna 12, E-50009, Zaragoza, Spain. E-mail: jlmedi@unizar.es

Abstract

The idea that mathematical ability and language are related in human cognition is an old one. It is commonly assumed in the philosophical tradition, in psychology and in cognitive science, generally implying that knowledge of numbers is indebted to knowledge of language. In this contribution I suggest a more specific model of the relationship between knowledge of numbers and knowledge of language in the light of developments in the neuroscience of numerical understanding and in linguistic theory. The proposed model places the evolutionary development of the syntactic component of the Faculty of Language at the basis of the transition from the innate sense of number that we share with other animals to the sophisticated mathematical ability of our species.

Keywords

arithmetic, number sense, language, syntax, numerals, successor function

1 Number Sense, Language and Arithmetic Capacity

In his influential essay *The Number Sense*, Dehaene (1997) argued that the numerical ability of human beings is based on a specific innate capacity which, contrary to what has been traditionally assumed, is not a by-product of language. Of course, this does not mean that language and numeracy are unrelated. What Dehaene proposes, in particular, is to distinguish between two dimensions of our species' numerical-mathematical knowledge: (i) an innate component (which we will call *the number sense*) and (ii) a



later cultural creation, apparently dependent on language, which implies the acquisition of a productive system of numbering in each individual (and which we can simply call *arithmetic capacity*).

The idea presented here, which does not contradict but rather complements Dehaene's view, is that the arithmetic capacity that goes beyond the innate number sense is not purely cultural, but is based on another natural "sense" that is different from the number sense. This capacity is one of the components of the so-called "language instinct" (Pinker, 1994), more specifically the syntactic or computational component of the language faculty, according to the model of language and its evolution by Noam Chomsky and collaborators (Berwick & Chomsky, 2016; Chomsky, 2020; Hauser et al., 2002).

The proposal I am about to make therefore implies that the arithmetic capacity of our species comprises three successive phases:

- i. The evolutionary development of number sense
- ii. The evolutionary development of the computational system of language (syntax)
- iii. The cultural development of number systems (linguistic and/or visual).

Only step (i) is independent of (and prior to) language, while steps (ii) and (iii) are language-dependent (and seem to be specifically human), albeit in different ways: step (ii) relies on the biological evolution of (one component of) the language faculty, while step (iii) requires the cultural innovation of speaking communities and their languages.

2 The Number Sense

Dehaene argues extensively that, contrary to conventional wisdom, human numeracy is not merely a cultural by-product of language or an incidental consequence of it, but is based on an innate system that we share with other animals and that can be found in pre-linguistic human infants. Indeed, *The Number Sense* is largely devoted to presenting evidence that animals such as pigeons, rats, dolphins or chimpanzees have an innate numerical capacity (the number sense) which is also present in human babies and which is therefore not based on the development of language.

In Dehaene's model, this innate number sense is conceived as a kind of cerebral analogue accumulator that allows the handling of small quantities without the need for counting. He even provides evidence that this capacity is associated with the inferior parietal cortex and that it is susceptible to selective lesions.

But, as Dehaene also points out, this innate sense of number does not explain the full development of the arithmetic capacity that humans exhibit, but is only the basis for it. The sense of number that we share with other species provides us with the very concept of numerosity and the direct or intuitive knowledge of the most elementary numbers, that is, the mental representations of 1, 2, 3 and (perhaps) 4. This ability allows us to

know instinctively, and without the need to count, that 1 is less than 2, that 2 is less than 3, and that 3 is less than any other quantity. But, to put it simply, from 3 or 4 onwards we have to count in order to conceive and manipulate quantities such as 13, 250 or 3,234. And counting and manipulating larger magnitudes, which go beyond this innate sense of number, is what is assumed to be dependent on language and cultural development in human beings.

3 A (Short) Digression: The Grammar of Number in Languages

Before proceeding further, it is interesting to note that, as Hurford observed in a commentary on Dehaene's book (Hurford, 2001), the number sense hypothesis may be supported by certain properties of the grammar of human languages.

The most notorious of these, and one that linguists observed early on, is that 3 (perhaps 4) is the upper limit for the inclusion of number in the grammar of languages. In fact, it is most common in the world's languages for grammatical number to be limited to the difference between singular and plural. Some languages, such as classical Greek, also have dual number (distinguishing 1, 2 and plural in morphology) and a few other languages have trial number (1, 2, 3 and plural). The existence of quadral in some languages has even been discussed (although both trial and quadral would be restricted to pronoun forms). What is important is that the processes of grammaticalisation of human languages show that speakers have been sensitive to these basic numbers (especially 1 and 2) since ancient times, and not to the larger ones.

Another interesting argument that Hurford mentions is that number names (numerals) for small numbers (up to 5) tend to have the same structure as simple adjectives (such as *blue* or *big*), as if these quantities were easily observable properties of sets of things, whereas larger numbers tend to have a more complex structure and seem to derive from strategies other than direct observation, for example because they are part of a memorised numbering sequence, or because they use syntactic composition (the comparison between *three* and *twenty-eight* makes the difference clear). This seems to show that our cognitive relationship to large numbers is different from our relationship to small numbers, in line with Dehaene's number sense hypothesis. As Hurford (2001) concludes, "It seems certain that in the historical development of languages, words for the lower numbers existed before words for higher numbers" (p. 69).

The comparative grammar of numerals provides further interesting data on the relationship between language and numbers. For example, in languages where numerals have gender, this is true only up to 4. If we look at case instead of gender, we find that the different case forms of numerals decrease as the denoted quantity increases (for example, as Hurford points out, in Russian the word for 1 has the six cases of that language, but the word for 10 is declined in only three cases). There are other formal

properties that distinguish the grammar of the numerals for 1–4 from the rest (see also Hurford, 1987), but they all point in the same direction, and are consistent with the experimental evidence provided by Dehaene: our cognitive relationship to the numbers provided by the number sense is different from the one we have with the numbers beyond the number sense on which our arithmetic capacity is based.

4 The Arithmetic Capacity

So it seems that the human mind is not designed for direct knowledge or mental representation of numbers greater than 3 or 4. In fact, Dehaene shows that our ability to mentally represent quantities decreases as we move up the number line from 1. It can therefore be concluded that the development of numerical efficiency above 3-4, i.e., that which supports the ability of most adults to know immediately that, for example, 189 is less than 212 (even though we cannot have mental representations of groups of 189 units or groups of 212 units and compare them), exceeds our natural numerical ability (number sense) and requires an artificial development of numerical systems, either with figures or digits (such as, for example, the Arabic system: 1, 2, 3, 21, 34, 86, 234, etc.), or with number words or numerals (as, for example, English: *one, two, three, twenty-one, thirty-four, eighty-six, two hundred and thirty-four*, etc.). Therefore, an innate number sense does not, by itself, explain our knowledge that 189 is smaller than 212 and 174 is smaller than 189, while 166 is smaller than all of them, nor does it explain our ability to create a (visual or verbal) representation of any conceivable natural number.

The next step in explaining this uniquely human capacity for arithmetic would therefore be to find out what our ability to create a productive system of numbers depends on. And this is where language comes in.

Contrary to what is usually assumed, however, the appeal to the possession of language, without further precision, is not in itself an explanation of this capacity. On the one hand, because many other species (chimpanzees, whales, birds, bees) have highly developed communicative capacities, which do not, however, make possible the arithmetic capacity demonstrated by our species. On the other hand, there are human cultures that have the capacity for language but have not developed symbolic or linguistic representations for large numbers, i.e. they lack the tools in their languages to record a number such as 2,834 (with figures) or *two thousand eight hundred and thirty-four* (with numerals). In such cultures, numbers larger than 3 or 4 are usually expressed in terms equivalent to *many* or *multitude* (see, e.g., Gordon, 2004; Pica et al., 2004).

So, if the mere possession of language as a communication system does not automatically explain arithmetic knowledge, the question we have to answer must be more specific: on what cognitive capacity does our ability to create a system adequate to express large quantities actually rest?

The short answer, which I will qualify later, is that this capacity is not language itself, but one of its components, syntax. Put simply, the ability to create an artificial number system (with figures or numerals) is actually based on the same ability that allows humans to use and understand an unlimited (potentially infinite) number of grammatical expressions in the languages they speak. In other words, arithmetic is based on the same thing as the compositional productivity of human languages, i.e. syntax. In this sense it is true that arithmetic depends on language, but it is not in fact a purely cultural construction, but is based on another cognitive endowment, which in this case seems to be exclusively human: the computational system we call syntax.

Gelman and Butterworth (2005) propose that advanced numerical knowledge is ontologically independent of language, but as I hope to show, once an adequate notion of language is considered, the hypothesis that arithmetic depends on the syntactic component of the language faculty is consistent with their arguments and simpler.

As far as we know, all languages have the capacity to create an unlimited number of grammatical and interpretable expressions. However, given that not all cultures (and languages) have developed a verbal or graphic number system (and certainly not all in the same way or at the same time), it seems clear that the specific development of number systems and arithmetic language is a cultural artefact.

Indeed, the grammatical regularity of different verbal number systems has been found to influence the rate of development of basic numerical skills, such as learning to count. Yang (2016a) derived an equation that predicts, with surprising accuracy, the number of irregular forms (for example, of verbs such as *broke*) that children learning their language can tolerate in order to induce the productive rule. Yang's equation relates the total number of verbs learned to the number of irregular verbs in order to predict how many regular verbs the child needs to learn in order to induce the correct productive rule. As far as we are concerned, this formula also explains why English-speaking children take longer to learn to count than Chinese-speaking children (Yang, 2016b). The reason for this is that in English there are 17 irregular forms in the first hundred numerals (i.e. 17 unpredictable forms that have to be learnt by heart, as the rest are formed by building on them, such as *twenty-nine*). Previous studies had shown a qualitative leap from 72, i.e. that English-speaking children learn to count in several slow, regressive stages, but that once they can count up to 72, there is no upper limit. Yang's equation, developed to explain the acquisition of the productive rule for forming the past tense of regular verbs, predicts that children will have doubts and problems counting until they reach exactly 72. However, in Chinese, only the first 12 digits need to be memorised, so according to Yang's equation, Chinese children will be able to count without limit once they have acquired 46 number words, which does indeed explain the previously observed higher precocity in counting ability of Chinese-speaking children compared to English-speaking children.

Of course, all children, whatever language they speak, eventually discover that it is always possible to add another number, a principle that apparently does not need to be learnt. The differences mentioned have to do with the regularity of the words and phrases (numerals) that have to be memorised in order to learn to count. Moreover, the degree of cultural development of number systems affects not only the precocity of learning to count, but also the ability to operate with numbers. But it is important to bear in mind that this cultural development could not have taken place without the support of a cognition ‘equipped’ with natural syntax, that is, with the principle of unlimited productivity that this capacity provides. Indeed, speakers of languages that do not have number systems are able to learn other languages that do, or to learn to ‘read’ Arabic numerical sequences with relatively little instruction (see reports in Gelman & Butterworth, 2005; Spelke, 2017; Yang, 2016b). The ability to know that you can always add another number to the last one you counted, although not part of the number sense, seems to be innate to all human beings, even if they do not use a number language or perform complex arithmetic operations. As Yang, points out, “cross-linguistic differences in numerical cognition are really just cross-linguistic differences in the numeral systems”, so that “the fact that some languages have a productive numeral system while other do not says nothing about the generative capacity of these languages nor the cognitive capacity of the speakers of these languages” (Yang, 2016b, p. 15).¹

In order to properly appreciate the sense in which I argue (following Chomsky’s insight) that syntax is the basis of arithmetic in human cognition, we first have two tasks: to properly understand what syntax is, and to consider what problem syntax had to solve in order to go beyond the innate number sense, i.e. why the mere allusion to language (without further concretisation) is ultimately insufficient to explain our unique arithmetic capacity. Let us begin with the second question and conclude with the characterisation of syntax relevant to my argument.

5 The Long and Winding Road to Numerical Language

The hesitant and convoluted cultural process towards greater numeracy, described in detail by Georges Ifrah (1994), has its starting point in the innate sense of number that we share with other species, and which is therefore independent of the faculty of language. Starting from this basic ability, the problem of expressing and communicating these ‘instinctive’ quantities involves associating them with (more or less arbitrary)

1) It is interesting to note that, although it would seem logical that number systems would develop more readily in cultures where they are useful (agricultural cultures or those with complex economies), there are also less developed cultures in this area (Pacific cultures such as Hawaiian, Tongan and Pukapukan) that have words for numbers up to 100,000 “without any obvious strictly practical use for them” (Hurford, 1987, p. 74).

words such as, for example, *one*, *two* and *three* in English, or with (more or less arbitrary) figures such as, for example, I, II and III in the Roman numeral system.

The greatest difficulty is in naming or writing numbers greater than 3 or 4. An early solution, still used in some cultures, was to establish a correspondence between numbers and parts of the body for counting or naming large numbers. As Ifrah reports, communities have been studied (such as certain Torres Strait Islanders) who counted visually by touching the fingers of the right hand one at a time, then the wrist (for 6), the forearm (for 7), and so on, until they reached the little toe of the opposite foot, reaching up to 40 or 50 units, depending on the different systems attested. From there, sticks or pebbles are often used to continue counting, according to Ifrah's description. In a gestural numeral system, when a number has to be counted or said with the hands occupied, the names of those body parts can be used as numerals, which explains why the etymology of numerals often leads to nouns for body parts. Nevertheless, even before the development of systematic notation or a developed numerical language, primitive systems such as these reflect an arithmetical cognitive capacity that goes beyond the innate number sense and seems to be based on the successor function (i.e. Peano's axiom that every number has a successor).

The best-known record of this journey from number sense to numeracy is, by definition, that of written number systems, i.e. the history of figures/digits. The starting point is the principle of correspondence between quantities and signs. For example, to record that we have seven sheep, we can collect seven pebbles (*calculi* in Latin), one for each sheep we see passing by, or, if we want a more stable representation, make seven marks on a stick or clay tablet: IIIIII. But such a sequence is difficult to read and interpret without counting, so it soon became practical to group the marks, as the Romans did, for example (IIII = V, then 7 = VII), which again shows that the representation of quantities greater than 3-4 is a challenge to our natural, limited ability to represent larger quantities.² Following the same principle on a different scale, the representation of 52 in Roman numerals will not be VVVVVVVVVVII, not even XXXXXII (once it is agreed that X is equivalent to "two hands"), but LII. In fact, in carved bones up to 30,000 years old (according to Ifrah, 1994), regular sequences of notches can be observed, grouped five by five, no doubt to save the person marking the register of collected pieces from having to count them one by one each time.

The use of figures to represent numbers dates back at least 5,000 years in Mesopotamia and Egypt, so it can be deduced that the numerals used to read such figures are at least as old (and probably much older). The oldest written figures are based on the additive principle, which means that the value of a numerical representation is the sum of the values of all the figures involved. For example, in Egyptian hieroglyphic

2) The use of V surely derives from the fact that we have five fingers on each hand, and fingers (*digiti* in Latin), as we all know, are the most primitive tool for calculating.

numeration, the number 3,273 was represented, according to Ifrah, by three times the thousands sign (three lotus flowers), twice the hundreds sign (two spirals), seven times the tens sign (seven inverted U's) and three times the unit sign (three vertical strokes), so that 15 digits would have to be used to represent this number. Although with greatly improved versions for abbreviation (including more different signs for intermediate units and the use of alphabetic characters), the ancient Greek, Roman, Jewish, Armenian and Arabic civilisations did remarkably well with systems based on this principle.

A crucial step along this path (bypassing countless intermediate and dead-end systems) is the development of a positional system, the essential key to which is the use of a small number of base numbers (typically 10) and their combination to represent larger quantities. The key is that in a positional system the same number, say 3, can represent three units, three tens, three hundreds, etc., depending on its position in the numerical representation. This system implies the ability to multiply as well as to add, so that, for example, the figure 327 represents that there are 3 hundreds, 2 tens and 7 units, i.e., it is the result of $3 \times 100 + 2 \times 10 + 7 \times 1$, as is clearly reflected in the internal syntax of its numerals in English: *three hundred and twenty-seven*. In fact, according to Ifrah, with the development of the multiplicative principle, the representation with figures and the verbal representation of numbers (which tends to be an 'oralisation' of positional sequences) began to coincide more closely.

The invention (or discovery) of 0 was crucial. It was originally used in India to replace a space when one of the quantities was missing. So 307 means that there are 3 hundreds, no tens and 7 units, and 307 (with a placeholder) reads much better than 3 7, which could well be confused with 37 (or even 3007). Since in a positional system the value of a number depends on its position in the sequence, 10 means that there is 1 ten and no units, from which 0 is also interpreted as denoting the null or empty set, another central tool in the development of modern mathematics. As Ifrah (1994) points out, the intellectual history of humanity is in a sense a long (and reverse) journey from 1 to 0.

The base-10 positional notation used throughout the world today is neither the only possible nor necessarily the best from a strictly mathematical point of view (there is binary notation), but it seems to follow from our anatomy, from the contingent fact that we have two hands. As Karl Popper (quoted by Dehaene, 1997, p. 117) pointed out, "natural numbers are the work of men, the product of language and of human thought".

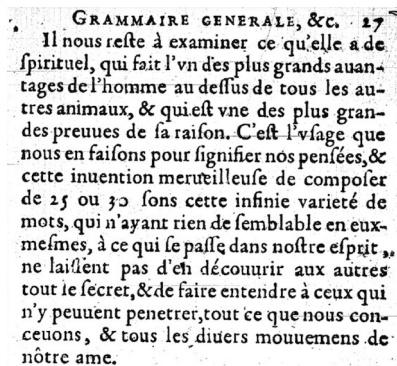
In any case, the fabulous invention of the positional notation system, including the 0 (which actually originated in India around the 6th century AD, even though they are called Arabic figures), implied a revolution in the calculating ability of successive generations. Note that the relationship between 5, 50 and 500 is transparent in this system, unlike the equivalent Roman numerals V, L, D (which is why the Romans, like many other peoples, did not do complex calculations with digits but with abacuses or counting tables). As Ifrah (1994, p. 777) points out, 'this principle seems so simple to us today that we forget that humanity stammered, hesitated and groped for millennia before

discovering it, and that cultures as advanced as the Greek and Egyptian civilisations completely ignored it’.

At this point, it is inevitable to relate the fascination aroused by the discovery of the possibility of representing any number, however large, with only 10 figures to the fascination expressed by Arnauld and Lancelot, authors of Port-Royal’s *Grammaire générale et raisonnée* of 1660, that a few sounds could express the infinite thoughts that a person is capable of conceiving. As they pointed out, ‘the marvellous invention whereby by using twenty-five or thirty sounds we can create an infinite variety of words’ that allow us to express ‘all that we can conceive and the most diverse movements of our soul’ is ‘one of the greatest advantages of human beings compared to other animals and one of the most conspicuous proofs of reason’ (my translation, see Figure 1).

Figure 1

Page 27 of the *Grammaire Générale et Raisonnée*



6 The Mystery of Discrete Infinity

In both cases, that is, in both number systems and languages, we are confronted with the same phenomenon: “the infinite use of finite means”, to use the words of Wilhelm von Humboldt (as quoted by Chomsky, 2019), or, as Chomsky often puts it, the property of “discrete infinity”. In Chomsky’s view (e.g. Chomsky, 1966), human language would be the only system, compared to the communication systems of other species, in which these two properties (infinity and discreteness) occur together, whereas the communication systems of other species are either discrete but limited, or unlimited but continuous.

But just as the positional number system we use today as the basis of arithmetic is a human cultural invention, the same cannot be said of the discrete infinity of human language. Certainly, writing is a human cultural invention, an invention parallel, in fact, to that of the figure systems we have summarised, with its same hesitations,

unrenounceable findings, and wide cultural diversity. In the vast majority of known traditions, the invention of the figure systems is indebted to the development of writing and, in particular, of alphabetic systems. The numeral system used by Pythagoras or Euclid was an alphanumeric system in which letters served as symbols for numbers, as well as representing in writing the phonemes of the spoken language.

However, contrary to what the authors of Port-Royal grammar suggested, the unlimited creative capacity of human language does not derive from the combinatorial character of phonology (roughly reflected in alphabetic writing), but from the combinatorial character of syntax, that is, from the capacity (apparently specific to human language) to create a potentially infinite number of distinct and semantically interpretable propositions from a necessarily limited number (as they are stored in memory) of simple concepts or notions.

Note that, then, the ability to count without limits and to conceive of infinitely large numbers cannot be explained as a mere consequence of the invention of compositional systems of representation (such as 2,345 or *two thousand three hundred and forty-five*), since the very invention of such systems of representation already presupposes the ability to conceive of a discrete infinity which is precisely what we are trying to explain.

Thus, the puzzle of the origin of arithmetic capacity is the same as the puzzle of the origin of human unlimited syntax, and the claim that the syntax of numbers derives from the syntax of languages, while relevant, leaves us without an answer to the central question. To put it another way, if the pressing question is how we explain the transition from a cognitive ability to represent only the numbers 1–4 to a cognitive ability to represent an infinite array of numbers, then appealing to the ability to create compositional numerals with a language (or with figures or digits) is not a satisfactory answer, because we still do not explain where the ability to create such compositional and recursive expressions in human languages themselves comes from.

The crucial idea, then, is that both human syntax and arithmetic share the same principle: the infinite use of finite means. Addressing this phenomenon is the central aim of the revolution in linguistic theory that led to the emergence of Noam Chomsky's generative grammar, which in turn was influenced by developments in mathematical logic and the theory of computability by authors such as Alan Turing, Emil Post and Alonzo Church (see Lobina, 2014).

7 Syntax as the Basic Property of Language

A central idea of Chomsky's linguistic theory is that a basic property of human language is that "each language constructs in the mind an infinite array of structured expressions each of which has a semantic interpretation that expresses a thought, each of which can be externalized in one or another motor system, typically sound, but not necessarily" (Chomsky, 2019, p. 17).

Note that this characterisation of the basic property of human language implies a conception of the language faculty as an integrated mosaic of at least three independent capacities (Hauser et al., 2002): (i) a conceptual-intentional system (responsible for the representation of meaning), (ii) a computational system (syntax in the strict sense, which creates complex compositional expressions by combining conceptual elements) and (iii) a sensorimotor system (responsible for translating these representations into muscular movements for interpersonal communication). The Chomskyan model argues that the uniqueness of language in our species derives from the evolutionary development of the computational system within a more complex cognitive organ whose other components (the conceptual-intentional and the sensorimotor systems) are widely shared with other species and have a much older and independent evolutionary trajectory.

More specifically, Chomsky conjectures that the relationship between the computational system or syntax and the other two systems is asymmetrical: the computational system would have been adaptively optimised for its interaction with the conceptual-intentional system and would have formed, together with it, an internal language of thought, while the relationship between this language of thought and the sensorimotor component (a necessary connection for the externalisation of thought and communication) would be secondary (and the main source of language diversity).³

In adopting this perspective, Chomsky and his followers are at odds with the widely held view that language evolved from ancestral systems of communication. But note that the hypothesis that syntax evolved in the service of an internal system of thought, rather than in the service of communication, has the fundamental advantage that it better explains the emergence of the most clearly distinctive property of human syntax, discrete infinity, which seems to be more useful for a system of thought than for a system of sense-signal matching. And this does not mean that there is no evolutionary continuity with ancestral communication systems in the other components of language.

Of course, the use of language for communication is essential to explain the development of human culture, and communication is a crucial phenomenon in human cognitive development. But this is not challenged by Chomsky's model, which simply states that the communication function does not explain why the central component of human language evolved, nor why it has the formal properties that it does. The evolution of a recursive (unbounded) syntax makes more adaptive sense if it was for the benefit of thought than if it was for the benefit of communication, since an unbounded syntax is not indispensable for efficient communication, but it is indispensable for the construction of new complex thoughts (which may be interesting to communicate).

3) "Parameterization and diversity, then, would be mostly –possibly entirely– restricted to externalization. That is pretty much what we seem to find: a computational system efficiently generating expressions interpretable at the semantic/pragmatic interface, with diversity resulting from complex and highly varied modes of externalization, which, furthermore, are readily susceptible to historical change" (Berwick & Chomsky, 2011, pp. 37–38).

Derek Bickerton (2014) called the puzzling fact that humans have far more intellectual capacity than they need to survive “Wallace’s problem”. Indeed, the co-discoverer of the theory of evolution by natural selection had to contradict himself when considering human cognition, as he could not imagine how the development of so-called “higher capacities” such as music, mathematics, art or language itself could have been adaptive for an archaic primate. But the problem is mitigated if the evolution of human language is seen as the development of a mosaic of relatively independent systems with different functions and evolutionary trajectories.

We are therefore in a position to conclude that the emergence of human syntax (understood as a recursive, unlimited computational system) had two relevant effects: it produced a creative language (from earlier, more limited systems of thought and communication) and, as we will show below, it could serve as a cognitive capacitation for arithmetic and the invention of number systems.

8 Elements of Syntax: Internal and External Merge

One of the most remarkable developments in syntactic theory in recent decades has been the demonstration that a large part of the rules, principles and constraints that need to be postulated in order to adequately characterise the syntactic structure of human languages can be derived from a minimal computational operation that Chomsky (1995) called *Merge*.

Although deriving all the structural complexity of the syntax of human languages from a minimal computational operation is controversial and empirically challenging (and runs the risk of reducing the descriptive adequacy of the theory), the explanatory advantage is clear: the emergence of a very simple capability, of minimal computational complexity, is more evolutionarily plausible than the emergence of a complex capability, and more compatible with the fact that such a capability cannot be very old in our evolutionary history (less than about 200,000 years, if specific to *Homo sapiens*, which leaves little room for a complex evolutionary process based on routine genetic mutation and natural selection).⁴ The simpler the principle or mechanism that must have arisen, the more plausible it is that it did so with minimal genetic variation (see Berwick & Chomsky, 2016).

The notion of Merge is really minimalist: the function Merge takes two linguistic objects, X and Y, and forms the unordered set {X, Y}; since it is a recursive operation (it reuses its own output), it can be applied again and merge Z with {X, Y}, creating

4) According to Miyagawa et al. (2025), genetic data appear to show that the modern faculty of language was already present in our species 135,000 years ago.

{Z {X, Y}}, and again, taking W, Merge can create {W{Z {X, Y}}}, and so on, indefinitely (limited only by external constraints of time, patience, memory capacity, etc.). Note that Merge does not create linear sequences, but hierarchical structures. There is considerable evidence that sentences in human languages are structured in this way, with words grouped into constituents, which in turn form larger constituents (for summarised arguments, see Everaert et al., 2015).

To give a less abstract, but very simplified example, we can imagine that an English speaker can merge the verb *ate* and the noun *carrots* to form *ate carrots* and then merge *Mary* to form *Mary ate carrots*. There is nothing to prevent this sentence from being part of another structure, so that we can merge the complementizer *that* and form *that Mary ate carrots* and then continue with the merge of the verb *said* to get *said that Mary ate carrots* and complete another sentence by merging *John* and end up with *John said that Mary ate carrots* (which in turn could be the continuation of *It seems that...*). Leaving aside many details that are relevant to the grammar of English (but irrelevant to our argument), what is important now is that this recursive use of Merge produces a hierarchical structure that the mind uses to assign a systematic semantic interpretation to the derivation (once the structure has been reconstructed from the written flat sequence or the continuous sound wave in the spoken language).

Chomsky calls the Merge operation we have considered *External Merge* because this process joins a linguistic object X to another linguistic object Y, which is different from X (and which is not already contained in X). As noted above, the External Merge operation has come to replace (and simplify) the systems of rules that were postulated in previous models for generating these structures (different phrase structure rules for sentences, for noun phrases, for verb phrases, etc.). The External Merge operation is the architect of the compositionality of linguistic expressions by combining the meanings of different elements and producing, so to speak, more complex meanings, such as argument and predicate structures, or propositions.

Traditional models have had to include, in addition to phrase structure rules, so-called “transformational rules”, because humans do not simply form sentences and pronounce them, but also modify sentences to change their interpretation, e.g. making interrogatives out of affirmatives or passives out of actives. The most striking effect of these structural transformations (which are common to all languages and involve changes in the information structure of utterances) is the “movement” of constituents, i.e. the appearance that certain fragments of a sentence appear to be interpreted in one place (e.g. where they were merged) but are pronounced in another (e.g. at the beginning of the main clause).

Let us look at a very simplified example. Our sentence *John said that Mary ate carrots* could be transformed into a partial interrogative if we replace *carrots* with a variable (*what*), assuming that this is what we want to ask about, and move it to the beginning: *What did John say that Mary ate?* To explain this transformation without resorting to a

complex set of ad hoc rules, we can stipulate that *what* has actually been merged in its natural place (object of *ate*), since we are in fact still interpreting *what* as the object of *ate*, even though it is at the beginning of the main clause.⁵

One way of explaining these cases, and of eliminating the cumbersome transformational rules, is to postulate that the so-called movement processes are in fact *Internal Merge* processes, i.e. processes in which an element already present in the derivation is copied and re-merged 'up'. Thus, we would have a starting point identical to the case of the affirmative, but with the variable *what* instead of *carrots*: *John said that Mary ate what* (derived from the External Merge of *what* with *ate*, of *ate what* with *Mary*, etc.). For reasons having to do with the semantic/pragmatic interpretation of the sentence, *what* is copied and merged back into the derivation (Internal Merge), resulting in *What John said that Mary ate what*.⁶ Although we cannot go into further detail, it is worth noting that Internal Merge builds a syntactic structure (*What John said that Mary ate what*) analogous to the Logical Form that actually has the interpretation of the interrogative sentence: 'for what X, it is true that John said that Mary ate X'.

9 Internal Merge and Arithmetic

So a single operation, Merge, could be behind the basic property of human language, discrete infinity. To understand how this ability is also behind arithmetic, we need to consider the difference between Internal Merge and External Merge. Chomsky (2020) argues that the computationally simplest operation is Internal Merge. His argument is based on the fact that External Merge has to access a 'dictionary', i.e. it has to successively select different objects (X, Y, Z, W, etc.) from a workspace which includes the entire lexicon, whereas Internal Merge feeds back, so to speak, repeating the merging of objects already present in the derivation, thus reducing the size of the workspace. This property of Internal Merge is the key to the emergence of arithmetic capacity.

As Chomsky argues, the explanation for the use of the (apparently more costly) External Merge operation in human language lies in its essential function (which is the creation of complex thoughts from simpler concepts), for if we had only Internal Merge, we would not have a language proper as a result, but only a language with a single 'word'. We would have the ability to combine a conceptual unit with itself, but not the relevant property of language to combine different conceptual units with each other:

5) One proof that this makes sense is that in languages where operators like *what* are accusative case when they are direct objects, they are also accusative when they are at the beginning.

6) Of course, nobody says that. In English, in these cases, the auxiliary appears 'above' its subject as a side-effect of the 'rise' of *what*, giving *What did John say that Mary ate what*, an expression which, if we do not pronounce the redundant second *what* (for economy of effort), is ready for delivery to the sensorimotor system: *What did John say that Mary ate?*

[W]hy does language ever use EM [External Merge]? The answer is straightforward. IM [Internal Merge] alone yields structures that have no interpretations at the CI [conceptual-intentional] interface. It does not yield expressions of thought, hence does not constitute an I-language [...]. There are no head-complement or XP-YP constructions, hence no theta structure. In fact, there can be only a single-membered lexicon. EM overcomes these inadequacies, yielding a possible I-language. Hence both EM and IM must be available: EM to yield an I-language in the first place, IM because it would require an unmotivated stipulation to bar the simplest operation. And indeed both are ubiquitous, with distinct semantic roles, a property sometimes called “duality of semantics”. (Chomsky, 2020, p. 10)

Chomsky’s approach suggests that the use of Internal Merge is a side effect of the need for External Merge in order to actually produce a language of thought. And Chomsky himself explicitly considers the possibility that, in addition to its use in language, Internal Merge is the basis of the successor function and thus of arithmetic:

Consider a system that only uses IM. It can be shown that this in effect yields the successor function, and with some limited tweaking, all of arithmetic. That is a suggestive conclusion, perhaps providing a solution to a problem that greatly troubled Darwin and Wallace: why is knowledge of arithmetic universal (so they assumed, correctly it seems, though like language and many other innately determined properties of the organism must be triggered by experience)? A serious problem for them, since it could not have evolved through natural selection. A possible answer is that Merge appeared at some point in the evolutionary record, perhaps along with *Homo Sapiens*, providing the Basic Property of language and also arithmetic. (Chomsky, 2020, p. 10)

If the adaptive feature for language was External Merge, then Internal Merge was a by-product of this process, which not only allowed for the “duality of semantics”, but also produced the ability to conceive of the successor function. The idea that Internal Merge, while not itself language-producing, could be the cognitive basis of arithmetic rests crucially on the fact that the iterative merging of a single ‘lexical unit’ would create the mental capacity to always add one more unit, limitlessly, just as we can always add one more word to any sentence. In an earlier paper, Chomsky (2008) already suggested that if we imagine a language with the simplest possible lexicon (only the ‘word’ *one*), the application of Internal Merge produces [*one[one]*], i.e. ‘two’, while a new application produces [*one[[one]one]*], i.e. ‘three’, and so on and so forth without formal limit. According to Chomsky, “Merge applied in this manner yields the successor

function. It is straightforward to define addition in terms of Merge (X, Y), and in familiar ways, the rest of arithmetic” (Chomsky, 2008, p. 139).⁷

The idea that Internal Merge produces the successor function is consistent with what Hurford (1987) calls “the conceptual/verbal hypothesis” to explain why there are no accidental gaps in the numerical lexicon. Drawing on Fodor’s (1976) reflections on the language of thought, Hurford suggests that one way to explain the strict sequence of natural numbers (there are no languages that have numbers for 7 but not for 6, etc.) is to assume that number words (e.g. *two* in English) are a kind of ‘label’ for language of thought phrases such as “the number after one”, and that only when that phrase has a verbal label (*two*) can the phrase “the number after two” be constructed (assuming that “the number after the number after one” is hard to process), so that “[g]iven such assumptions about how concepts for numbers can become present to the mind, and given an initial state in which the concepts of number, 1, and successor are possessed, along with the words for them, it follows that concepts for particular numbers can only become available to the mind in strict succession, and must be accompanied ‘one step behind’, as it were, by the invention of number words” (Hurford, 1987, p. 96).

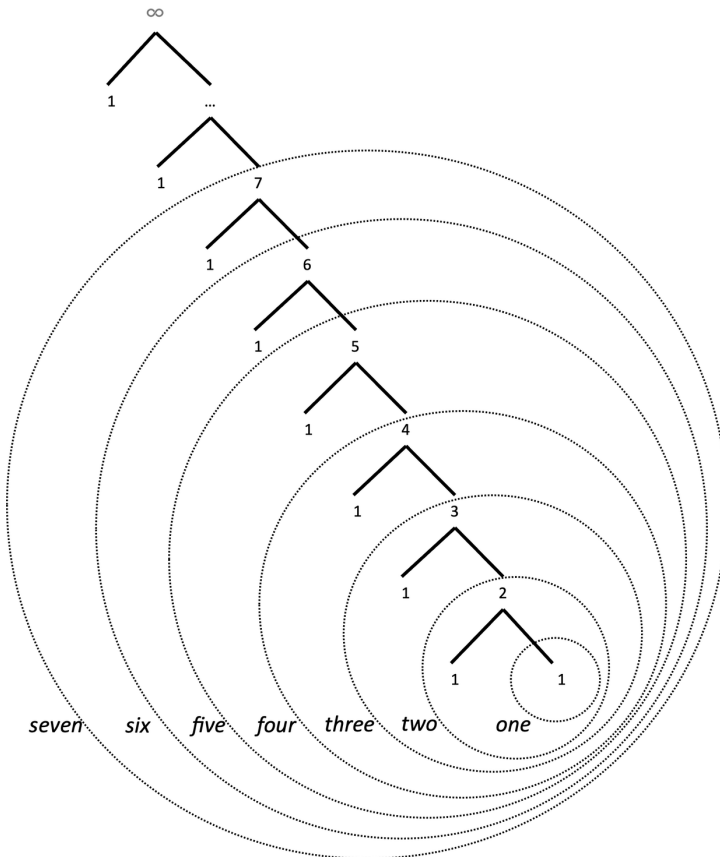
The hypothesis that Internal Merge (whatever its role in language) is necessary for External Merge provides an elegant explanation for the ability to induce the successor function, once it is applied to the innate number sense. Figure 2 shows schematically how the repeated internal merge of ‘concept’ 1 (unit) can generate the numerical sequence and serve as the conceptual basis for the uneven and varied lexicalisation of this conceptual numerical structure in the different languages of the world (of those that have number systems). Of course, each system has its historical peculiarities, and the influence of the invention of the positional numeral systems we have considered is evident in many of the current systems of numerals, but this does not mean that arithmetic capacity can be explained directly as a result of the “invention” of numerals.⁸

7) An anonymous reviewer objects to this approach, saying that “if we adopt Chomsky’s notion of Merge, there should be a label which stems from one of the merged items.” But, in fact, Merge is independent of the labelling operations that are required to account for headedness. The introduction of Merge separates three independent factors that have been conflated in previous syntactic theories: compositionality, labelling, and ordering. The key to understanding the cognitive roots of the successor function is compositionality, not labelling or order, which are not part of Merge.

8) Hurford objects to explanations based on the innateness of Peano’s axiom that the words for 2 never reflect that they were conceived as the sum of *one plus one* (Hurford, 1987, p. 100 ff.). However, his empirical observations would be explicable if we consider that the number sense already provides 2 as a basic unit, so that the successor function is only necessary for counting (and enumerating) from 2 (or 3) onwards. There is also the possibility that the successor function, once incorporated into human cognition, is ‘retroactively’ applied to intuitive basic numbers, although this is not reflected in the terms used to name them.

Figure 2

Recursive Internal Merge of Number 1 and Lexicalisation With English Numerals



The diagram in Figure 2 does not imply that I am claiming that the conceptual representation of, say, the number 23 (*twenty-three*) involves twenty-two mergers of 1 every time that word is used or every time that number is operated on, since in this particular case an External Merge of *twenty* and *three* is clearly observed (see Hurford 1987 for a detailed study of the logical basis and historical development of complex numerals in the world's languages).⁹

9) Although Hurford's pragmatic explanation of the syntactic structure of complex numerals is relevant, it seems that the parallel cultural development of numerical systems, especially the base-10 positional system, also has a direct influence on this syntax (e.g. reading 385 as *three hundred eighty-five*).

Note that even if it is true that the terms for the basic numbers originate from memorised sequences of arbitrary words (as is usual for small numbers, usually up to 10), the cardinality principle (which implies that the word we have stopped at in an enumeration corresponds to the quantity of objects in a collection) can only be explained if we consider a conceptual scheme such as that in Figure 2, in which obtaining the next number results from a Merge operation, although this is not reflected in the syntactic or morphological structure of the numeral. Thus, contrary to what Hurford (1987, p. 305) concludes, the cardinality principle would not be, as he suggests, the only innate aspect of numerical capacity that has nothing to do with language, but would follow directly from the linguistic operation Internal Merge.

In strictly logical terms we know that $7 = 1+1+1+1+1+1$ is just as true as $7 = 6+1$ (since $6 = 1+1+1+1+1$), but the second equation ($7 = 6+1$, based on $6 = 5+1$, etc.) is much more psychologically realistic, and corresponds to the “meaning” of 7 that follows directly from the derivation formed by Merge. This is due to a central property of Merge: structure preservation. As mentioned above, the Merge operations that form sentences do not form linear chains of words (weak generation), but rather preserve the structure of the derivation (strong generation).¹⁰

An anonymous reviewer questions whether the development of modern language coincides with advanced mathematical skills, on the grounds that even finite numbers such as 12345678901234567890 are unpronounceable in any number system of any language. This clearly shows the reverse interpretation of the model I have proposed, which is from number to numeral and not from numeral to number. But the existence of numbers that are impossible to pronounce is in fact parallel to the existence of grammatical sentences that are impossible to process: it is not a problem of knowledge, but of processing. We may not be able to pronounce the number 12345678901234567890 using numerals, but we can affirm, without being mathematicians, that 12345678901234567890 is the result of adding ‘one’ to 12345678901234567889, probably because Internal Merge is part of our cognitive endowment.

10 Conclusion

It is suggestive to think that the cognitive source of the concept of unity necessary to derive the successor function from Internal Merge is precisely the sense of number given to us by our evolutionary ancestors, and that the same evolutionary innovation that gave us syntax, gave us arithmetic.

¹⁰ As Lasnik and Uriagereka (2022) point out, in the study of language we need a computational system that generates not only strings of symbols, but also “the structure assigned to those strings”, because “in human language the most basic distinctions are captured only if close attention is paid not just to the output string, but also to the process that took us there” (Lasnik & Uriagereka, 2022, p. 4).

The interpretation of the relationship between syntax and the successor function proposed here is different from (but, I believe, not incompatible with) that offered by Yang (2016b). According to Yang, it is the acquisition of the productive syntax of numerical language that enables the induction of the successor function. The same trajectory “from numeral to number” is advocated by Hurford (1987). However, a productive linguistic number system implies the cognitive presence of External Merge, which in turn implies the cognitive presence of Internal Merge, so that the relativistic effects Yang aptly points out have more to do with the acquisition of the specific numerical system of each language (essential for the development of arithmetic capacity) than with the true induction of the successor function.

The previous discussion also allows us to reaffirm the conclusion that Hurford draws from his own consideration of the relationship between numbers and language: “The capacity to reason about particular numbers, above about 3, comes to humans only with language” (Hurford, 1987, p. 306). But the approach I have presented suggests a deeper connection between language and numbers, so that we can better qualify this conclusion without contradicting it.

And furthermore, the present approach is also consistent with the evidence presented by Gelman and Butterworth (2005) that the concept of numerosity is independent of language (in the sense in which they use the term *language* as equivalent to numerals or number words), since in the present model, although the successor function follows from language (in a more technical sense of the term *language* than is usual in psychology and neuroscience), I do not assume that the successor function is a consequence of learning the numerals specific to each language, since numerals, as shown in Figure 2, belong to the externalisation of language and its processing, but they do not create the numerical concepts. In that sense, my approach coincides with Gelman and Butterworth in that “we need to distinguish possession of the concept of numerosity itself (knowing that any set has a numerosity that can be determined by enumeration) from the possession of representations (in language) of particular numerosities” (2005, p. 6), although it remains true that arithmetic capacity depends on a central property of language (i.e. Internal Merge). To use John Locke’s statement that they mention (Gelman & Butterworth, 2005, p. 8), “concepts of numbers are independent of their names”, but not, I would add, of their syntax.

Therefore, we can conclude that our mathematical capacity derives from (at least) two natural instincts: the number sense (which provides at least the notion of 1) and the syntactic capacity (unbounded Merge) which, beyond the number sense, provides the successor function. These two natural capacities are, of course, the necessary cognitive basis for the cultural development of the sophisticated number systems that underlie the rest of mathematics, science and technology.

In this sense, this approach also gives a new meaning to the suggestive words of Georges Ifrah (1994, p. 30), according to which, ‘figures bear witness, more and better

than the Babel of languages, to the profound unity of human culture', since the profound unity of culture (in figures/digits and also in languages) is only possible thanks to the unity of human nature.

Funding: This research was supported by a grant from the Government of Aragon (Spain) to the Psylex Group (H11_23R).

Acknowledgments: I would like to thank two anonymous reviewers for their helpful comments.

Competing Interests: The author has declared that no competing interests exist.

References

- Berwick, R. C., & Chomsky, N. (2011). The Biolinguistic program: The current state of its development. In A. M. Di Sciullo & C. Boeckx (Eds.), *The biolinguistic enterprise* (pp. 19–41). Oxford University Press.
- Berwick, R. C., & Chomsky, N. (2016). *Why only us: Language and evolution*. The MIT Press.
<https://doi.org/10.7551/mitpress/9780262034241.001.0001>
- Bickerton, D. (2014). *More than nature needs. Language, mind, and evolution*. Harvard University Press.
- Chomsky, N. (1966). *Cartesian linguistics*. Harper & Row.
- Chomsky, N. (1995). *The minimalist program*. The MIT Press.
- Chomsky, N. (2008). On phases. In R. Freidin, C.P. Otero & M. L. Zubizarreta (Eds.), *Foundational issues in linguistic theory* (pp. 133–166). The MIT Press.
- Chomsky, N. (2019). *The UCLA Lectures*.
<https://www.its.caltech.edu/~matilde/ChomskyUCLALectures.pdf>
- Chomsky, N. (2020). Fundamental operations of language: Reflections on optimal design. *Cadernos de Linguística*, 1(1), 1–13. <https://doi.org/10.25189/2675-4916.2020.v1.n1.id271>
- Dehaene, S. (1997). *The number sense. How the mind creates mathematics*. Oxford University Press.
- Everaert, M. B. H., Huybregts, M. A. C., Chomsky, N., Berwick, R. C., & Bolhuis, J. J. (2015). Structures, not strings: Linguistics as part of the cognitive sciences. *Trends in Cognitive Sciences*, 19(12), 729–743. <https://doi.org/10.1016/j.tics.2015.09.008>
- Fodor, J. A. (1976). *The language of thought*. The Harvester Press.
- Gelman, R., & Butterworth, B. (2005). Number and language: How are they related? *Trends in Cognitive Sciences*, 9(1), 6–10. <https://doi.org/10.1016/j.tics.2004.11.004>
- Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science*, 306(5695), 496–499. <https://doi.org/10.1126/science.1094492>
- Hauser, M., Chomsky, N., & Fitch, T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science*, 298(5598), 1569–1579. <https://doi.org/10.1126/science.298.5598.1569>
- Hurford, J. R. (1987). *Language and number: The emergence of a cognitive system*. Basil Blackwell.

- Hurford, J. R. (2001). Languages treat 1–4 specially. *Mind & Language*, 16(1), 69–75.
<https://doi.org/10.1111/1468-0017.00157>
- Ifrah, G. (1994). *Histoire universelle des chiffres* (2 Vols.). Robert Laffont.
- Lasnik, H., & Uriagereka, J. (2022). *Structure: Concepts, consequences, interactions*. The MIT Press.
<https://doi.org/10.7551/mitpress/11276.001.0001>
- Lobina, D. J. (2014). What linguists are talking about when talking about... *Language Sciences*, 45, 56–70. <https://doi.org/10.1016/j.langsci.2014.05.006>
- Miyagawa, S., DeSalle, R., Nóbrega, V. A., Nitschke, R., Okumura, M., & Tattersall, I. (2025). Linguistic capacity was present in the *Homo sapiens* population 135 thousand years ago. *Frontiers in Psychology*, 16, Article 1503900. <https://doi.org/10.3389/fpsyg.2025.1503900>
- Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science*, 306(5695), 499–503. <https://doi.org/10.1126/science.1102085>
- Pinker, S. (1994). *The language instinct. How the mind creates language*. The MIT Press.
- Spelke, E. S. (2017). Core knowledge, language, and number. *Language Learning and Development*, 13(2), 147–170. <https://doi.org/10.1080/15475441.2016.1263572>
- Yang, C. (2016a). *The price of linguistic productivity: How children learn to break the rules of language*. The MIT Press.
- Yang, C. (2016b). *The linguistic origin of the next number* [Unpublished manuscript].
<https://ling.auf.net/lingbuzz/003824>