# Cultural Bootstrapping of Quantity Terminologies Based on Approximate Number Sense

Nina Gierasimczuk
Dariusz Kalociński
Franciszek Rakowski
Jakub Uszyński

**Abstract**

The language coordination problem is the question of how a population of agents can spontaneously develop a shared lexicon usable in communication. We present an agent-based model evolving quantity expressions, i.e., expressions that refer to numerocities. Our model is based on the existing approach to simulate the emergence of color expressions. We substitute its perceptual layer by cognitive model of approximate number sense (ANS). Simulations result in communicatively usable quantity expressions which proves that the model solves the language coordination problem for the domain of quantities. Finally, we look at two widely-discussed semantic universals—convexity and monotonicity. We express them within our model and determine how well they are represented in the emerging quantity terminologies. We also compare these results with the degrees of convexity and monotonicity of languages that emerge without employing approximate number sense to see which properties stem from the nature of perception and which have their source in other layers of the model.

## 1 Introduction

Even though natural languages seem quite different from each other, linguists observed many striking similarities. It is widely believed that languages share some non-trivial properties, known as *linguistic universals*. They have been hypothesised at many levels of linguistic analysis, including phonology (Hyman Larry M., 2008), syntax (Chomsky, 1965; Croft, 1990), and semantics (Barwise and Cooper, 1981).

Once a linguistic universal is proclaimed, the question arises as to why the property in question is so ubiquitous. The most common answers relate universality to human cognition (see, e.g., Christiansen and Chater, 2016). Hence, the features and constraints of the human cognitive system that might become relevant during learning and using a language have the potential to affect the linguistic structure and possibly result in a universal.

Universals have been linked to various aspects of language, e.g, communication (Kemp and Regier, 2012), complexity (Szymanik and Thorne, 2017; Rolando et al., 2018), and learnability (Steinert-Threlkeld and Szymanik, 2020). In this paper we look at how the shared subsystem of the human cognition responsible for the perception of numerocities can shape the fragment of language denoting quantities, possibly leading to linguistic universals.

Human cognition is hypothesized to be equipped with an evolutionarily old mechanism of number cognition called the *approximate number sense* (ANS, Dehaene, 1997). ANS allows instantly perceiving quantities at the cost of accuracy, with an error proportional to the perceived input. While it is not surprising that ANS, being a general purpose mechanism, is recruited during linguistic tasks involving number perception such as quantifier verification, a more intricate issue is whether the ANS has the potential to shape the semantic structure of the fragments of languages for denoting quantities (henceforth, *quantity terminologies*). To answer this question we need a model that solves the language coordination problem, i.e., the problem of how a population of individuals comes up with a language. For this purpose, we use one of the state-of-the-art solutions to this problem, originally applied to color terminologies (Steels and Belpaeme, 2005). We take the original model and substitute entirely its perceptual layer by the ANS, leaving the other layers intact. The model is described in Section 2. Section 3 defines notions of *semantic* and *pragmatic* meaning within the model and introduces two semantic universals that can be attributed to quantity terms, namely *convexity* (Gärdenfors, 2000) and *monotonicity* (Barwise and Cooper, 1981; Peters and Westerståhl, 2006). Section 4 constitutes the experimental part where we contrast the degrees of convexity and monotonicity of quantity terminologies resulting from the ANS-based model with the respective degrees obtained from a model where the number perception of agents is far more accurate, with perception error being independent of the perceived input.

# 2 Model

The model consists of three modules. The first is environment which provides stimuli to individual agents. The second module describes perception, i.e. how agents cognitively represent incoming stimuli and how they can differentiate between them using their cognitive representations. The third module, a linguistic one, introduces a cultural dimension where two agents interact, trying to inform each other about some relevant aspect of a shared context by emitting a signal. We describe each module in succession.

## 2.1 Environment

In the traditional semantic literature, a quantifier is understood as a class of situations (technically speaking—models) which the quantifier can be truthfully applied to (Peters and Westerståhl, 2006). For example, the quantifier denoted by *some* is a class of models $(A, B)$ such that $A \cap B \neq \emptyset$ (*Some boys are rude* is true if, in given situation, the set of boys intersected with the set of rude children is non-empty). For simplicity, we omit the layer of models and deal directly with quantities (henceforth, stimuli) which correspond to certain quantitative properties of models.

In the simplest case, a stimulus is a positive integer. This type of stimulus is referred to as numeric-based. In our experiments, a numeric-based stimulus is taken from the range $\{1, 2, \ldots, 20\}$. It is convenient to think of a numeric-based stimulus as the size of a set of objects presented to an agent. Another type of stimulus that we use is slightly more involved. Instead of having two positive integers, we have two: $n$ and $k$. The relevant property of the two integers is their ratio $\frac{n}{k}$. This type of stimulus is referred to as quotient-based. In experiments, we restrict the range of quotient-based stimuli in the following way. A denominator is taken from the set $\{1, 2, \ldots, 100\}$ and it is combined with a non-greater numerator, giving the desired fraction. Again, although our model does not include any representation of sets or any set-based structures, it might be helpful to have in mind how quotient-based stimuli are related to real-word situations. A raw situation might be thought of as a finite set of objects $U$ (a universe) together with a subset $A \subseteq U$ (a property). Such a situation is usually denoted by $(U, A)$. To give an intuitive example, consider $U$ to be a set of students in a classroom and $A$—a set of girls in the classroom. The corresponding integers are $|U| = n$ and $|A| = k$. The quotient-based stimulus $\frac{k}{n}$ thus represents a ratio of girls in the classroom.

Each time agents interact with each other they are presented with a context. In the model of Steels and Belpaeme (2005), a context is a small set of color stimuli. Here, a context is a set of two stimuli, either numeric-based or quotient-based, depending which type of stimulus is set to be used in a simulation. Hence, for numeric-based and quotient-based stimuli, contexts are of the form $\{n_1, n_2\}$ and $\{\frac{n_1}{k_1}, \frac{n_2}{k_2}\}$, respectively (with restrictions mentioned in the previous paragraph).

Another important feature of the environment is how stimuli are actually presented to agents. We resort to a simple pseudo-random mechanism. If we need a numeric-based stimulus, an integer is drawn from the uniform discrete probability distribution $\mathcal{U}\{1, 100\}$. If we need a quotient-based stimulus, we first draw a denominator $k \sim \{1, 100\}$ and then draw a numerator $n \sim \mathcal{U}\{1, k\}$.

## 2.2 Perception

Agents' perception consists of two levels. First, an agent forms basic responses to stimuli in the form of so-called reactive units (Section 2.2.1). Second, an agent forms higher-level representations, so-called categories, which are built on top of reactive units (Section 2.2.2). The discrimination process, formalized as the discrimination game, recruits these two levels to differentiate between stimuli and suitably adapts the level of categories in case of discriminative success or failure (Section 2.2.3).

### 2.2.1 Reactive units

An internal representation of a stimulus is referred to as **reactive unit**. Reactive units may be viewed as cognitive tools for representing simple external properties objects. In the Steels model, reactive units are gaussian functions defined on the CIELAB color space: each reactive unit is centered on a certain point (so-called focal point) in the color space where it has its maximum and fades proportionally to the distance from its center. Crucially, the standard deviation of all reactive units is kept constant.[1]

Transition from colors to quantities is not immediate. Luckily enough, number perception in humans and other species is a well-researched topic. One of the well-known theories is the theory of approximate number sense (henceforth, ANS) (Dehaene, 1997). ANS is a fast and evolutionary old mechanism for rough representation of quantities. Our considerations are based on the following model of ANS. Given a number $n$, its ANS representation, denoted by $R_n$, is given by:

$$R_n \sim \mathcal{N}(n, \gamma n), \tag{1}$$

where $\mathcal{N}(n, \gamma n)$ is the normal distribution with mean $n$ and standard deviation $\gamma n$. We assume $\gamma = 0.1$ throughout the paper.

There are a few obvious but important things to keep in mind about the ANS model defined above. First, numbers are represented only approximately via random variables: agent-wise, there is always some uncertainty regarding the exact value causing a given perception. Second, probabilities behave normally which implies—among other things—that the numbers further from the mean appear less likely to be causes of a given sensation. Third, the standard deviation of the ANS representation $R_n$ is directly proportional to $n$. This implies that the greater the number, the wider the scope of numbers that are just easily confusable with it. The latter property is a realization of the Weber law in the cognitive domain of quantities (Fechner, 1966). A few examples of reactive units for numeric-based stimuli are presented in Figure 1.

Reactive units are based on ANS representations of numbers defined above. In case of a numeric-based stimulus $n$, a reactive unit representing it is simply $R_n$. In practice, $R_n$ is represented only approximately so that the actual probability density function associated with $R_n$ is not supported on the entire real line but rather on a sufficiently long line segment which encompasses almost 100% of the true probability mass. The approximation of $R_n$ is denoted by $\widetilde{R}_n$.

Reactive units for quotient-based stimuli are also based on ANS representations. Consider a quotient-based stimulus $\frac{n}{k}$. The ANS representation of $\frac{n}{k}$ is the ratio distribution $R_{n/k}$ given by

$$R_{n/k} \sim R_n / R_k. \tag{2}$$

[1] It is worth noting that the CIELAB space is an antropocentric color space in the sense that equal distances between colors in that space correspond to approximately equal perception of color distance by humans. Hence, keeping the constant variance parameter in the definition of guassian functions defined on the CIELAB space is justified.

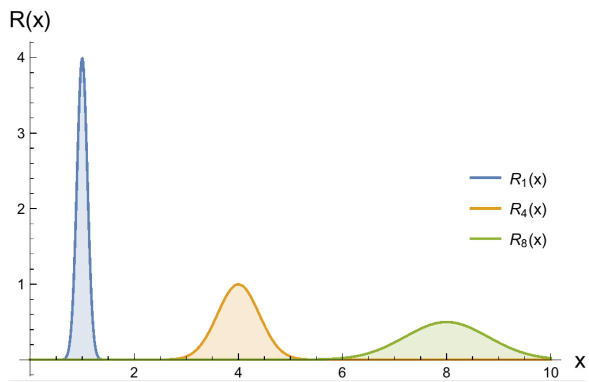


Figure 1. Examples of reactive units for numeric-based stimuli $1, 4$ and $8$. Reactive units become more extensive for greater quantities.

Since $R_n, R_k$ are gaussian distributions, we could, in principle, compute the probability density function of $R_{n/k}$ based on known analytical expressions (see, e.g., (Hinkley, 1969; Pham-Gia et al., 2006)). In practice, we resort to an empirical approximation of $R_{n/k}$. We first draw two equally-sized large samples $(n_i)_{i=1}^s$ and $(k_i)_{i=1}^s$ from $R_n$ and $R_k$, respectively. Then we derive a density histogram based on ratios $(\frac{n_i}{k_i})_{i=1}^s$. Finally, we obtain a desired approximation by interpolating the resulting histogram. The approximation of $R_{n/k}$ is denoted by $\widetilde{R}_{n/k}$. A few examples of reactive units for quotient-based stimuli are presented in Figure 2. When the exact form of a reactive unit is irrelevant, we denote it

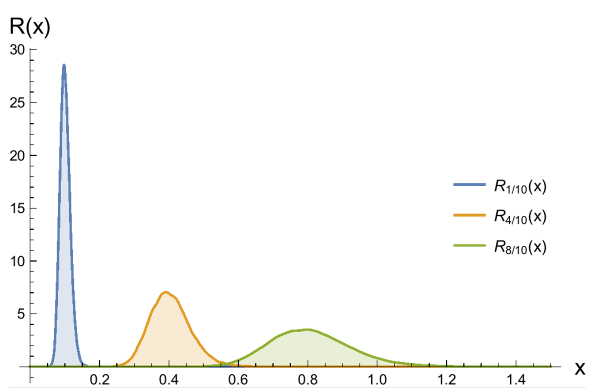


Figure 2. Examples of reactive units for quotient-based stimuli $1/10, 4/10$ and $8/10$. The reactive unit $R_{8/10}$ is visibly different from a normal distribution because of its longer right tail.

by $z$, possibly with subscripts. We write $z(x)$ we mean $p_z(x)$, i.e. the value of the probability density function associated with $z$ for $x \in \mathbb{R}$.

While simple numerical expressions have been empirically studied with respect to their ANS representations, quotients have not. We decided to follow here a possible mathematically justified interpretation of quotients. We leave the cognitive validation of this proposal to further work.

#### 2.2.2 Categories

Categories are built on top of reactive units. They form a more abstract perceptual layer that groups reactive units into more general coherent entities. These entities induce a coarse partitioning of the perceptual space so that an agent can make meaningful distinctions between stimuli. This intuitive description will perhaps become more meaningful when technicalities are revealed.

Each category consists of a finite number of reactive units $z_1, z_2, \ldots, z_k$ and corresponding non-negative weights $w_1, w_2, \ldots, w_k$. Upon reception of a stimulus, a category becomes activated. The magnitude of this activation is governed by a linear combination of reactive units and weights:

$$c(x) = \Sigma_{i=1}^{k} w_i z_i(x). \tag{3}$$

This approach is the same as in (Steels and Belpaeme, 2005).

The most important function of a category is to provide responses to stimuli. Roughly speaking, a response of a category $c$ to a stimulus $s$, denoted by $resp(c, s)$, indicates how much the stimulus seems to belong to the category. In (Steels and Belpaeme, 2005), a response of a category $c$ to a color stimulus $\overline{x}$ from the CIELAB space is identified with its activation $c(\overline{x})$. Here, response of a category $c$ to a stimulus (either numeric- or quotient-based) represented by a reactive unit $z$ is given by:

$$< c|z >= \int_{-\infty}^{+\infty} c(x)z(x)dx. \tag{4}$$

The rationale behind this definition is twofold. First, the stimulus is not perceived directly but through reactive unit so we cannot assume that the response of $c$ to a stimulus $x$ is simply $c(x)$.[2] Second, intuitively, the response of $c$ to a stimulus represented by $z$ should be proportional to the joint contributions of $c$ and $z$ over the same regions of the space of stimuli. For example, if the activation of $c$ and $z$ for a given region is high, the response should be higher, accordingly. Similarly, if the activation of $z$ is low for a given region, the response should also be low. These intuitive properties are realized by the scalar product of $c$ and $z$ shown in equation (4).

#### 2.2.3 Discrimination Game

A discrimination game is an individual cognitive mechanism the main function of which is to, on an input of a stimulus, output a category which distinguishes

[2] This is justified in the original Steels model because complex activation patterns of the nervous system, produced in response to colour stimuli, are reduced to three-dimensional points in the CIELAB space.

a that stimuli (the topic) from other stimuli in the context (Steels, 1997). The precise sense of this distinction is explained below. In the case of color perception, a context is a finite set of color stimuli (Steels and Belpaeme, 2005). Here, contexts consist either of numeric- or quotient-based stimuli. As Pauw and Hilferty (2012) we restrict contexts to two stimuli.

**Discrimination game**

1. An agent is presented with a context consisting of two stimuli $q, q'$ one of which is considered to be the topic.
2. If the agent has no categories, the game fails. Let $c_1, c_2, \ldots, c_k$ be the categories of the agent. Let $r_i$ be the response of $c_i$ to $q$ and $r'_i$ the response of $c_i$ to $q'$, for $i = 1, 2, \ldots, k$ (responses are computed according to normal or extreme response, see Section 2.2.2). If all $r_i$s or all $r'_i$s are 0, the game fails.
3. Let $j = \operatorname*{argmax}_i r_i$ and $j' = \operatorname*{argmax}_i r'_i$. In words, for each stimulus, we choose (an index of) a category that generates the strongest responsethe best-matching categories for $q$ and $q'$, respectively.
4. The discrimination game is successful if $j \neq j'$. Upon success, one of the best-matching categories, either $c_j$ or $c_{j'}$ is declared as the winning category, depending on whether $q$ or $q'$ is the topic. If $j = j'$, the discrimination game fails.

**Success** When a discrimination game is successful, the winning category is reinforced so that if it were reused in the same game, its response for the topic would be even stronger. In general, reinforcement is to enhance chances of winning categories to win similar games in the future. Let $q$ be the topic and $c$ a winning category. Let $w_i, z_i,\ i = 1, 2, \ldots, k$, be the weights and reactive units of $c$, respectively. Reinforcement is computed according to the following expression:

$$w_i = w_i + \beta \cdot resp(z_i, q). \tag{5}$$

The parameter $\beta$ is called the learning rate. In our simulations we assume $\beta = 0.1$. The expression (5) is rather self-explanatory: weights of the winning category $c$ are increased proportionally to the corresponding reactive units' responses to the topic.

**Failure** When a discrimination game fails, an agent tries to repair the system of categories to avoid a similar malfunction in the future. Let $C$ be the set of categories of the agent who has just played and unsuccessful discrimination game for the context $q, q'$.

1. Suppose the agent has no categories, i.e. $C = \emptyset$. Then a new category $c$ is created, consisting of one reactive unit $z_1$ representing the topic, with an associated weight $w_1 = 0.5$. After repair, we have $C = \{c\}$.

2. Suppose all $r_i$s or all $r_i'$s are 0 (see the second step of the discrimination game). It means that and agent has a gap in her system of categories as no category gets activated for $s$ or $s'$, respectively. In that case, a lacking category is added in the same way as in the first point above (for $s$, if $s$ does not yield any positive response, otherwise for $s'$).

3. Suppose that the agent has successfully passed two steps of the discrimination game and $j = j'$. If the discriminative success of the agent is lower than the discriminative threshold $\delta$, the agent creates a new category $c$ consisting of a single reactive unit representing the topic, with initial weight 0.5. If the discriminative success of the agent is greater than or equal to the discriminative threshold $\delta$, the category $c_j$ (which is the same as $c_{j'}$) is adapted by appending a new reactive unit representing the topic, with initial weight 0.5. We assume $\delta = 90\%$.

After discrimination game and reinforcement/repair, the weights of all reactive units of all categories of the agent are decreased by a small factor $0 < \alpha < 1$ (we assume $\alpha = 0.1$ in simulations):

$$w_i = \alpha w_i. \tag{6}$$

This results in slow forgetting of categories. If all weights of a category are smaller than 0.01, the category is deleted from the repertoire of categories of the agent.

## 2.3 Language

Language provides a bridge between words from agent's dictionary and external stimuli via categories. An important feature of words is that they can be directly transferred between agents while categories and reactive units cannot. Dictionary is a dynamic object. Initially, it is empty but can grow and shrink during the lifetime of an agent. The role of this feature will become evident in Section 2.3.1 where we describe the guessing game. Technically, words in the dictionary are short strings (we use the package `gibberish` to generate nonsensical pronounceable random strings, see Haskins, 2016).

Let $F$ be a dictionary and $C$ a set of categories. Language is a function $L : F \times C \to [0, \infty)$ which assigns a non-negative weight $L(f, c)$ to each pair $(f, c) \in F \times C$. $L(f, c)$ indicates the strength of the connection between word the $f$ and the category $c$. If $L$ is a language, we denote its dictionary by $F(L)$ and its set of categories by $C(L)$. Language is a dynamical object that changes during the lifetime of an agent. These changes are caused both by additions and deletions performed on the dictionary and categories, and by changes in the connections between words and categories in the language.

### 2.3.1 Guessing Game

A guessing game is a basic language game between two agents, the speaker and the hearer, both facing a common context consisting of a finite number of stimuli (Steels, 1997). The speaker wants to turn hearer's attention to one stimulus in the context, called the topic. Crucially, only the speaker is aware which stimulus constitutes the topic. The goal of the speaker is to utter a word so that the hearer, upon hearing it, will be able to guess the topic correctly.

Let $D_S, C_S, L_S$ be the dictionary, the categories and the language of the speaker. Similarly, we have $D_H, C_H, L_H$ for the hearer.

1. The speaker and the hearer are presented with a common context consisting of two stimuli: $q_1, q_2$. One of the stimuli, denote it by $q_S$, is the topic. The information whether $q_S = q_1$ or $q_S = q_2$ is available only to the speaker.

2. The speaker tries to discriminate the topic $q_S$ from the context by playing the discrimination game (see Section 2.2.3). If the winning category is found (denote it by $c_S$) the game continues. Otherwise the game fails.

3. The speaker searches for words $f$ in $D_S$ which are associated with $c_S$, i.e. such that $L_S(f, c_S) > 0$.

   If no associated words are found (i.e., $L_S(f, c^s) = 0$ for all $f \in D_S$) or $L_S$ is empty, the speaker creates a new word $f$ (i.e., $D_S := D_S \cup \{f\}$), sets $L_S(f, c^S) = 0.5$ and and conveys $f$ to the hearer.

   Now suppose that some associated words are found. Let $f_1, \ldots, f_n$ be all words associated with $c_S$. The speaker chooses $f$ from $f_1, \ldots, f_n$ such that $L_S(f, c^S) \geq L_S(f_i, c^S)$, for $i = 1, 2, \ldots, n$, and conveys $f$ to the hearer. The choice of one of the words is consistent with the concept of pragmatic meaning.

4. The hearer looks up $f$ in her lexicon $D_H$.

   If $f \notin D_H$, the game fails and the topic is revealed to the hearer. The repair mechanism is as follows. First, the hearer adds $f$ to her lexicon (i.e., $D_H := D_H \cup \{f\}$). Next, the hearer plays the discrimination game to see whether or not she has a category capable of discriminating the topic. If one is found, say $c$, the hearer creates an association between $f$ and $c$ with the initial strength of 0.5 (i.e., $L_H(f, c) = 0.5$).

   Suppose the hearer finds $f$ in $D_H$. Let $c_1, c_2, \ldots, c_k$ be the list of all categories associated with $f$ in $L_H$ (i.e., $L_H(f, c_i) > 0$ for $i = 1, 2, \ldots, k$). The hearer chooses $c_H$ from $c_1, \ldots, c_k$ such that $L(f, c_H) \geq L(f, c_i)$ for $i = 1, 2, \ldots, k$. The hearer points to the stimulus, denoted by $q_H$, that generates the highest response for $c_H$ (i.e., $q_H = \underset{q \in \{q_1, q_2\}}{\operatorname{argmax}}\, resp(c_H, q)$).

5. The topic is revealed to the hearer. If $q_S = q_H$, i.e., the topic is the same as the topic guessed the hearer, the game is successful. Otherwise, the game fails.

**Success** The speaker increases the strength between $c^S$ and $f$ by $\delta_{inc}$ and decreases associations between $c_S$ and other word forms by $\delta_{inh}$. The hearer increases the strength between $c_H$ and $f$ by $\delta_{inc}$ and decreases the strength of competing words with the same category by $\delta_{inh}$.

**Failure** The speaker decreases the strength between $c_S$ and $f$ by $\delta_{dec}$. The hearer decreases the strength between $c_H$ and $f$ by $\delta_{dec}$.

We assume $\delta_{inc} = \delta_{inh} = \delta_{dec} = 0.1$.

# 3 Convexity and Monotonicity

Let us first define the notion of *meaning* in our model. Roughly speaking, the meaning of a word is the set of categories associated with it. While defining the general notion of meaning we do not care about contextual factors and other competing words the meaning of which might occupy more or less similar areas of the perceptual space. However, we can also think of a more restrictive concept of the *pragmatic meaning* of a word, which is defined by looking also at the contribution of contextual factors and interactions with other words. Pragmatic meaning will then consist of stimuli for which a speaker is most likely to utter precisely the word in question.

Let us begin with our notion of **meaning**. Let $f$ be a word in language $L$ (see Section 2). We denote the meaning of $f$ in $L$ by $[[f]]^L_s$, but we also write $[[f]]_s$ if the underlying language is known. The meaning is a subset of the space of stimuli $Y = \{y_1, y_2, \ldots, y_n\}$ (recall that stimuli are either positive integers or rationals and that we restrict the space of stimuli to a finite set). We provide a formal definition of the meaning of a word below but the main idea behind it can be easily stated in plain words: a stimulus $y$ contributes to the meaning of $f$ if at least one category $c$ associated with $f$ generates a positive response to $y$.

$$[[f]]^L_s =_{df} \{y \in Y : \Sigma_{c \in C(L)} L(f, c) \langle c | \widetilde{R}_y \rangle > 0\}. \tag{7}$$

A justification for this definition comes from the analysis of the guessing game. Consider a situation in which there is a context consisting of two stimuli $y, y'$ with $y$ as the topic. The speaker utters a word $f$. Now, the hearer searches for the category with the highest association to $f$ and finds $c$. To choose $y$ and win the game, it suffices that $c$ produces a higher response to $y$ than to $y'$. Note that, in principle, any stimulus $y$ generating a positive response to $c$ can outperform (in terms of the response magnitude) some other stimulus $y'$. Hence, the interpretative process on the part of the hearer suggests that $y$ indeed contributes to the meaning of $f$.

Before we proceed to the definition of pragmatic meaning, we want to make a remark about our approach with regard to the famous distinction between intension and extension (Frege, 1892). First, note that the meaning of a word is identified with its extension—it is a subset of the set of possible quantity stimuli. However, the meaning of a word is obtained via its intension—a fragment of the weighted categorical network that supports all referential and truth-value judgements. One might even view this internal representation of language from the perspective of algorithmic theory of meaning (Tichy, 1969). In fact, all the processes in speaker and hearer that lead to uttering a word and guessing the topic can be (and are) represented by programs. Finally, let us stress that the (pragmatic) meaning is defined agent-wise. What is usually construed as the extension of a word, namely an entity that is shared by the members of community, is not given right away but is rather gradually constructed in the laborious process of repeated guessing games leading to various repairs which make individual languages slowly converge to a sufficiently shared common language.

Now we turn to the notion of **pragmatic meaning** which is slightly more involved.

$$[[f]]_p^L = \{y \in Y : \exists\, c\, [c \in \operatorname*{argmax}_{\substack{c \in C(L) \\ \langle c|\widetilde{R}_y\rangle > 0}} \langle c|\widetilde{R}_y\rangle \ \wedge\ f \in \operatorname*{argmax}_{\substack{f' \in F(L) \\ L(f',c) > 0}} L(f',c)]\}. \tag{8}$$

In words, $y$ is an element of the pragmatic meaning of $f$ if $f$ is maximally (and non-negatively) associated with a category $c$ that generates a maximal (and non-negative) response to $y$. The contextual dependence of a pragmatic meaning is realized by selecting a category $c$ that generates a maximal (and non-negative) response to $y$. From the usage perspective, such a category may be considered as the most salient one for $y$. The interaction of $f$ with the rest of the lexicon is taken into account by checking whether $f$ maximizes the association strength with $c$ compared to other words.

## 3.1 Convexity

The mathematical notion of convexity is easily supported by our geometric intuitions. A subset of points of the real line is convex if it constitutes a line segment. For the two-dimensional Euclidean space, a set of points is convex if it intersects every line into one line segment. In general, the notion of convexity is defined for any vector space and can be extended to functions defined over vector spaces (Rudin, 1966).

The notion of convexity has paved its way into cognitive research thanks to the theory of conceptual spaces (Gärdenfors, 2000, 2014). According to this theory, meanings can always be represented geometrically, and natural categories must be convex regions in the associated conceptual space. For evidence supporting this theory, see, e.g., (Jäger, 2010), and especially (Chemla et al., 2019), where the notion of convexity is investigated both theoretically and experimentally for logical words such as quantifiers.

Convexity of a set in a conceptual space is usually defined using an in-between relation that should satisfy several constraints (Gärdenfors, 2000). However, as has been observed by Chemla et al. (2019), one might depart from a strictly ordered set for which a natural in-between relation (defined by: $b$ is between $a$ and $c$, in symbols $a \textcircled{b} c$ if $a < b < c$ or $c < b < a$) satisfies the required axioms of in-betweenness. In our case, a strictly ordered set is the set of stimuli $(Y, <)$, where $Y$ consists of either natural numbers or rationals and $<$ is a *less than* relation. A definition of convexity by Chemla et al. (2019) states that $S \subseteq Y$ is **convex** if for any $a, b, c \in S$: $a\,b\,c$ implies that $b \in S$. Hence, the meaning of $f$ in $L$ (be it pragmatic or semantic) is convex if it constitutes a convex set in the space of stimuli $Y$. The above definition of convexity of meaning is sufficient for our purposes and can be derived from a more subtle analysis of convexity for quantifiers presented by Chemla et al. (2019).

## 3.2 Monotonicity

Natural languages realize only a very limited subset of quantifier meanings as denotations of simple determiners (see, e.g., Keenan and Stavi, 1986). It has been argued that all such naturally occurring quantifiers posses certain well-behaved properties. This has led researchers to put forward a list of quantifier universals (Peters and Westerståhl, 2006). One such a universal is *monotonicity*. Consider the following sentences:

Many good chess players know advanced tactics. (9)

Many good chess players know tactics. (10)

Note that (10) follows from (9). The only difference in the above sentences is that *advanced tactics* is replaced by a more inclusive expression *tactics*. The inference would be still valid if we replaced any expression $A$ with a more general expression $B$. This shows that the validity of the above inference is entirely based on the meaning of the quantifier *many*. The property of meaning that guarantees that is called (upward) monotonicity. In simple words, a quantifier is upward monotone if whenever it can be truly applied to a set of objects it can be truly applied to any of its supersets.

Another variant of monotonicity is called downward monotonicity. Consider the following sentences:

Few beginning chess players know tactics. (11)

Few beginning chess players know advanced tactics. (12)

Again, (12) follows from (11) and the only difference is that *tactics* is substituted by a more specific term *advanced tactics*. The inference would be still valid if we replaced any expression $B$ with a more specific expression $A$.

A quantifier is called monotone if it is either upward or downward monotone. The claim concerning the universality of natural language quantifiers states that all simple determiners denote monotone quantifiers (Barwise and Cooper, 1981; Peters and Westerståhl, 2006).

We shall attribute monotonicity to meanings of words in our model. The justification of this choice is rather obvious: a pragmatic meaning represents only the fragment of the overall meaning—the one that is most contextually and linguistically salient. For example, even though *most* can be truly used in situations where all objects posses a given property (e.g., *Most aardvarks are insectivores*), it may be more salient to use *all* instead.

Let us try to pinpoint the notion of monotonicity from the perspective of the model described in Section 2. Let $f$ be a word in language $L$ and let $Y$ be a space of quantity stimuli. We say that $[[f]]_s^L \subseteq Y$ is **monotone** if it is upward closed with respect to $\leq$ (i.e., if $y \in [[f]]_s^L$ and $y \leq y'$ then $y' \in [[f]]_s^L$) or downward closed with respect to $\leq$ (if $' \in [[f]]_s^L$ and $y \leq y'$ then $y \in [[f]]_s^L$).

# 4 Experiments

The code for performing experiments and data visualization has been written in Python and Mathematica and is freely available at `anonymized_url`. All experiments are based on the same parameters the values of which are indicated earlier in the text. A simulation consists of 30 separate trials. Within each trial, a population of 10 agents evolves quantity terminologies across 3000 successive steps. At each step agents are paired randomly to play a guessing game. We have performed separate simulations for numeric-based and quotient-based stimuli. In addition, we have varied whether agents are equipped with the ANS representation or not. Overall, this gives us four conditions.

In the first experiment we show that the model defined in Section 2 is valid, i.e., that it provides a solution to the language coordination problem in the domain of quantities (Section 4.1). Next we define the notion of convexity within our model and check the extent to which this property is represented in languages emerging with and without the approximate number sense (Section 4.2). We provide an analogous definition and experiments for monotonicity (Section 4.3).

## 4.1 Model validity

We proceed to determining whether the model defined in the previous section solves the language coordination problem in the domain of quantities. We use three types of metrics to evaluate the dynamics of the model: discriminative success, communicative success, and number of active terms in the lexicon. Similar measures have been used in previous studies (see, e.g., Steels and Belpaeme, 2005; Pauw and Hilferty, 2012). Details about metrics are presented below, followed by the presentation of results and discussion.

Discriminative success describes how well agents cope with the problem of distinguishing topic from the rest of the context. Let us define the measure of discriminative success more formally. Let $ds_j^a = 1$ if the $j$-th discrimination game of an agent $a$ is successful. Otherwise, let $ds_j^a = 0$. The cumulatinve discriminative success of an agent $a$ at game $j$ for the last $n$ games is given by:

$$DS(n)_j^a = \frac{\Sigma_{i=max(0,j-n)}^{j} ds_i^a}{n}. \tag{13}$$

The cumulative success of a population $A = \{a_1, a_2, \ldots, a_m\}$ for the last $n$ games is defined as:

$$DS(n)_j = \frac{\Sigma_{a \in A} DS_j^a}{m} \tag{14}$$

In practice we always check the measures of success for the last fifty games ($n = 50$).

The measure of communicative success reflects how well a population performs in guessing games. Intuitively, this measure indicates the percentage of successful linguistic interactions with other agents. The minimal requirement regarding the communicative success is that at some point it will grow above 50% and stay above that threshold later on. Notice that the communicative success will never be greater than the discriminative success since successful communication entails successful discrimination. Hence, the minimal requirement for the discriminative success is that it should attain 50% level. Lower levels are not be enough to bootstrap a communicatively useful language.

To define the concept of communicative success more formally, let $cs^a_j = 1$ if the $j$-th guessing game of an agent $a$ is successful. Otherwise, let $cs^a_j = 0$. The definitions of cumulative communicative success $CS(n)^a_j$ of an agent $a$, and of cumulative communicative success $CS(n)^A_j$ of a population $A$, at game $j$ for the last $n$ guessing games are defined in the same way as $DS(n)^a_j$ and $DS(n)^A_j$, respectively (replace $DS$ with $CS$ and $ds$ with $cs$).

Finally, we also track the number of active words in the lexicon. We call a word $f \in L$ active if $[[f]]^L_p \neq \emptyset$, i.e., if the pragmatic meaning of $f$ is not empty. The active lexicon of an agent is the set of her active words. The notion of active lexicon represents the actually used fragment of the entire language of an agent.

### 4.1.1 Results

Figure 3 shows the results. We have plotted the discriminative success and the

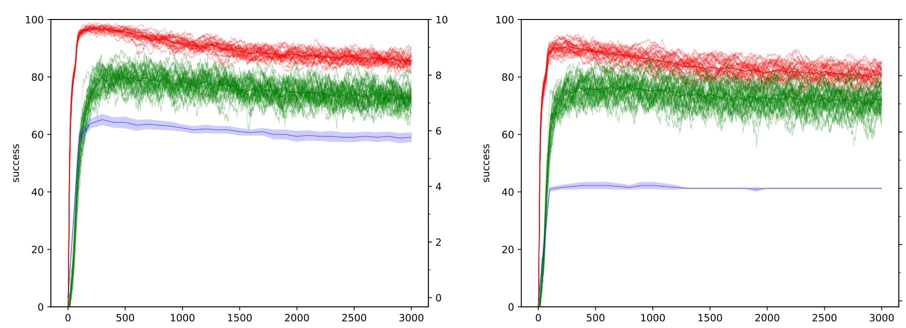


Figure 3. Discriminative success (red lines), communicative success (green lines), and size of active lexicon (blue lines, with scale on the right side of the plot). The left-hand side plot shows numeric-based stimuli, quotient-based are on the right.

communicative success for each trial. Each red line represents the discriminative success of a population evolving its language within a single trial. Each green line represents the corresponding communicative success. The blue line represents the average size of active lexicons of agents across trials.

The discriminative success attains acceptable levels, mostly above 80%. This shows that the categorical networks of the agents allow them to successfully discriminate the topic from the rest of the context in the vast majority of discrimination games. The communicative success lags behind the discriminative success, varying between 60% and 90%, with an average between 70% and 80% which is strictly above the minimal usability success rate of 50%. This shows that the evolving languages are communicatively useful, allowing the interlocutors to successfully communicate about the topic in the majority of guessing games. The languages evolving for numeric-based stimuli have lexicons with 6 active terms on average. The active lexicons for quotient-based stimuli have active lexicons of mean size 4.[3]

The results seem to justify the conclusion that the model defined in Section 2 solves the language coordination problem for the domain of quantities. Assessing the validity of the model is necessary if any further conclusions are to be drawn from it.

With that kind of evidence in our hands, this model may serve as a hypothetical mechanistic theory explaining the transition from the stage of having no quantity terminology to the stage in which such a terminology exists. This allows us to turn our attention to a different kind of question, namely: does the theory predict certain properties of real quantity terminologies?

## 4.2 Convexity

The convexity of a language for a single agent is calculated as the number of active words that have convex pragmatic meanings, divided by the size of the active lexicon. The population-level convexity is the average convexity of languages of all agents in the population. The plots in Figure 4 show population-level convexity averaged across all 30 trials. We observed that convexity appears

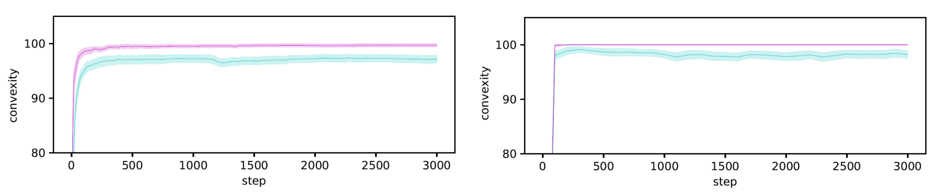


Figure 4. Population-level convexity, with ANS (magenta lines) and without ANS (cyan lines). The left-hand side plot shows numeric-based stimuli, quotient-based are on the right.

naturally with high frequency in all conditions, but also that ANS representation

[3]As a side note, we mention that both success metrics attain higher levels for the model without ANS. This should not be surprising—turning ANS off leads to more precise number perception and the evolution of more fine-grained terminologies.

facilitates convexity to the larger extent. We can hypothesize that if one is attached to intuition that cognitive concepts should be convex in the appropriate conceptual space, it is more plausible to assume vague meanings of words.

### 4.3 Monotonicity

The monotonicity of a language for a single agent is calculated as the number of monotone meanings corresponding to words from the active lexicon, divided by the size of the active lexicon. The population-level monotonicity is given by the average monotonicity of languages of all agents in the population. The plots in Figure 5 represent population-level monotonicity averaged across all 30 trials. In this case, there is no doubt that ANS contributes significantly to

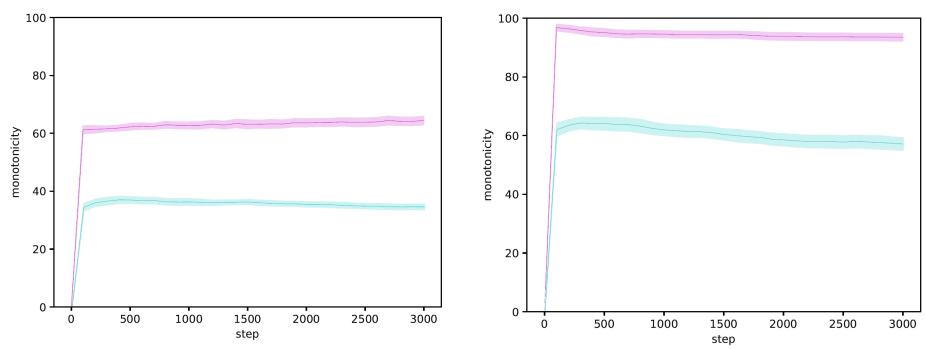


Figure 5. Plots for the numeric-based and quotient-based stimuli are to the left and to the right, respectively. Condition without ANS is given in cyan, while the ANS condition in magenta.

the emergence of monotonicity in language. Again, speculatively speaking, the more relaxed, vague semantics of quantity expressions under ANS likely allows an easier upward or downward merge with other categories.

## 5 Conclusions

We presented an agent-based model evolving quantity expressions (quantifiers). Experimenting with quantity expressions in agent-based simulation is a recent topic (for different kind of coordination experiment with quantifiers see Kalociński et al., 2018). Our model in the present paper is based on the existing approach to simulate the emergence of color expressions (Steels and Belpaeme, 2005). The perceptual layer of our model is inspired by the cognitive model of approximate number sense (ANS). Agent-based simulations based on agent interaction in so-called guessing games result in communicatively usable quantity expressions. This proves that the model solves the language coordination problem for the domain of quantities. Finally, we look at two widely-discussed semantic universals: convexity and monotonicity. We express them in our model and determine that they are well represented in the emerging quantity terminologies. ANS turns out to facilitate both properties, but the positive effect is especially evident in the case of monotonicity.

To get a better understanding of the resulting quantity terminologies let us look at Figure 6, which shows 5 agents taking part in the same evolution process over 300 trials (the timeline is represented by the $x$ axis). The $y$ axis gives the size of the stimuli (we focus here on the numerical ones). Each color represents a different word, whose semantics spans over a range of stimuli. We see words appearing and disappearing with time (according to the rules of discrimination and guessing games). We can also observe that after the initial variation the separate terminologies start resembling each other structurally, following the ANS-based perception—there are several words with restrictive fine-grained meaning, for small quantities, for larger ones the words have broader meanings. All emerging terminologies are convex—at step 300 there are no words that contain 'holes' in the depicted quantity interval between 1 and 20. This image illustrates the fact that even though the languages of individual agents differ among each other, the satisfactory level of communication is still possible to achieve. The common cognitive apparatus and the pressure to cooperate in communication probably contribute to that effect.

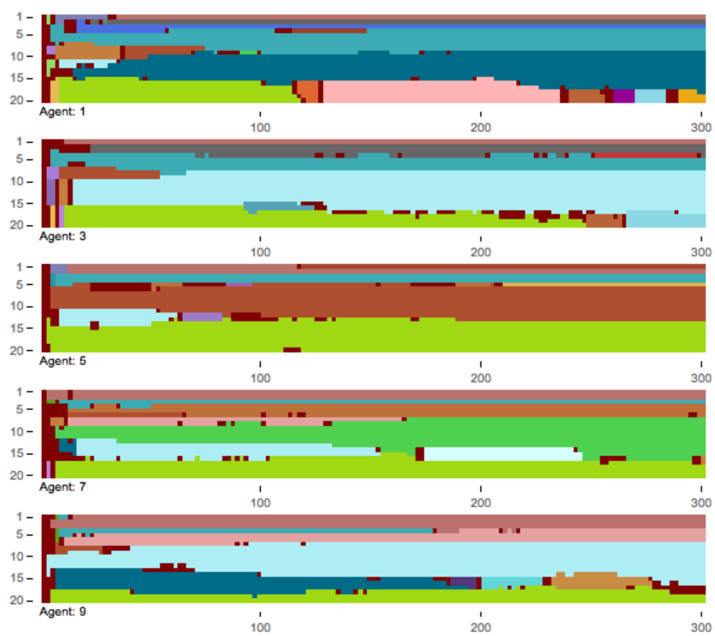


Figure 6. The evolution of individual agents' quantity terminologies over 300 trials.

## 6 Acknowledgement

This research is supported by the Polish National Science Centre grant OPUS 2015/19/B/HS1/03292.

## References

Barwise, J. and Cooper, R. (1981). Generalized Quantifiers and Natural Language. *Linguistics and Philosophy*, 4(2):159–219.

Chemla, E., Buccola, B., and Dautriche, I. (2019). Connecting Content and Logical Words. *Journal of Semantics*.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press.

Christiansen, M. H. and Chater, N. (2016). *Creating Language: Integrating Evolution, Acquisition, and Processing*. MIT Press.

Croft, W. (1990). *Typology and Universals*. Cambridge University Press, Cambridge.

Dehaene, S. (1997). *The number sense: How the mind creates mathematics*. OUP USA.

Fechner, G. (1966). *Elements of psychophysics. Vol. I.* Elements of psychophysics. Vol. I. New York, Holt, Rinehart and Winston.

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100:25–50. siehe auch Wikipedia:Über Sinn und Bedeutung.

Gärdenfors, P. (2000). *Conceptual spaces: The geometry of thought*. MIT press.

Gärdenfors, P. (2014). *The geometry of meaning: Semantics based on conceptual spaces*. MIT Press.

Haskins, G. (2016). Gibberish: a pseuro-word generator. `https://pypi.org/project/gibberish`.

Hinkley, D. V. (1969). On the ratio of two correlated normal random variables. *Biometrika*, 56(3):635–639.

Hyman Larry M. (2008). Universals in phonology. *The Linguistic Review*, 25(1-2):83.

Jäger, G. (2010). Natural Color Categories Are Convex Sets. In Aloni, M., Bastiaanse, H., de Jager, T., and Schulz, K., editors, *Logic, Language and Meaning*, pages 11–20, Berlin, Heidelberg. Springer Berlin Heidelberg.

Kalociński, D., Mostowski, M., and Gierasimczuk, N. (2018). Interactive Semantic Alignment Model: Social Influence and Local Transmission Bottleneck. *Journal of Logic, Language and Information*, 27(3):225–253.

Keenan, E. L. and Stavi, J. (1986). A semantic characterization of natural language determiners. *Linguistics and Philosophy*, 9(3):253–326.

Kemp, C. and Regier, T. (2012). Kinship Categories Across Languages Reflect General Communicative Principles. *Science*, 336(6084):1049–1054.

Pauw, S. and Hilferty, J. (2012). The emergence of quantifiers. In *Experiments in Cultural Language Evolution*, volume 3, pages 277–304.

Peters, S. and Westerståhl, D. (2006). *Quantifiers in Language and Logic*. Oxford University Press, Oxford.

Pham-Gia, T., Turkkan, N., and Marchand, E. (2006). Density of the Ratio of Two Normal Random Variables and Applications. *Communications in Statistics - Theory and Methods*, 35(9):1569–1591.

Rolando, S., Kirby, S., and Smith, K. (2018). Compressible kinship terminologies are more learnable than less compressible alternatives. In Cuskley, C., Flaherty, M., Little, H., McCrohon, L., Ravignani, A., and Verhoef, T., editors, *The Evolution of Language: Proceedings of the 12th International Conference (EVOLANGXII)*. NCU Press.

Rudin, W. (1966). *Real and Complex Analysis*. McGraw-Hill series in higher mathematics. McGraw-Hill.

Steels, L. (1997). Constructing and sharing perceptual distinctions. In van Someren, M. and Widmer, G., editors, *Machine Learning: ECML-97: 9th European Conference on Machine Learning Prague, Czech Republic, April 23–25, 1997 Proceedings*, pages 4–13. Springer Berlin Heidelberg, Berlin, Heidelberg.

Steels, L. and Belpaeme, T. (2005). Coordinating perceptually grounded categories through language: A case study for colour. *Behavioral and brain sciences*, 28(4):469–489.

Steinert-Threlkeld, S. and Szymanik, J. (2020). Ease of learning explains semantic universals. *Cognition*, 195:104076.

Szymanik, J. and Thorne, C. (2017). Exploring the relation between semantic complexity and quantifier distribution in large corpora. *Language Sciences*, 60:80–93.

Tichy, P. (1969). Intension in terms of Turing machines. *Studia Logica*, 24(1):7–21.