Inferring semantic representations underlying the meanings of numerals

Introduction What semantic representations underlie the meanings of numerals? Let us assume that numerals denote either (i) numbers, when they non-ambiguously pick out a number — for instance, the denotation of English *two* is in (1) (we leave aside the question of whether this results from truth-conditional meaning or from pragmatic enrichment, cf. Spector 2013); or (ii) sets of numbers, when their meaning is 'more than n', for some number n (cf. Table 1).

(1) $[\![\text{two}]\!] = 2$ (2) $[\![\text{two}]\!] = 1 + 1$ (3) $[\![\text{two}]\!] = s(1)$

(2) and (3) are denotationally equivalent to (1), with s the successor function. However, does our semantic representation of *two* involve 2 as a semantic primitive, or $1 + 1$, or $s(1)$? As semantic representations cannot be observed, they need to be inferred. Multiple approaches to this challenge have been developed, inferring semantic representations from behavioural data (Hackl 2009, Pietroski et al. 2009, Lidz et al. 2011, Piantadosi et al. 2012, 2016, Knowlton et al. 2021) or from typological generalizations (Züfle and Katzir 2021). We put forward a novel approach to inferring semantic representations using data on the optimality of the languages' simplicity/informativeness trade-off. We use numerals as a case study, building on the simplicity/informativeness trade-off analysis by Xu et al. (2020). Importantly, the approach can be applied to any domain for which cross-linguistic semantic data is available.

Hypotheses In this project, we focus on numerals denoting numbers or sets of numbers 1–10. We follow a tradition in semantics and philosophy of language to think about semantic representations in terms of combinations of primitive concepts (Fodor 1975, Pietroski 2018), and assume that the semantic representations of numerals are composed from a certain set of primitive number concepts PRIM, functions $+$, $-$ and successor s ($s(n) = n + 1$), and relation 'greater than' $>$ ($> n = \{x \in \mathbb{N} | x > n\}$) (cf. also Xu et al. 2020). Semantic primitives and operations may have different complexities (e.g., $+$ may be semantically more complex than $>$). This can be modeled as *weight* w_x for a primitive or operation x. Our research question is what PRIM and weights underlie semantic representations of numerals. We explore ten hypotheses, according to which PRIM contains $[1, \ldots, n]$, with $n \in \{1, \ldots, 10\}$. For each of these hypotheses, we consider all possible assignments of two values (1 and 2) to w_{PRIM} (the weight assigned to elements of PRIM), w_+, w_-, w_s, $w_>$. This amounts to $2^5 \times 10 = 320$ hypotheses.

Method Natural languages differ in terms of how complex they are to represent and in terms of how informative they are (i.e. how precise a communication they allow for). For instance, focusing on numbers 1-10, some languages have numerals for only a few of them, while others have numerals for each of them (cf. Table 1) – the former are simpler, but the latter are more informative. Simplicity and informativeness are in a tension: languages cannot both be maximally simple and maximally informative. This tension is known as the *simplicity/informativeness trade-off problem*. There can be many *optimal solutions* to this problem: the set of optimal solutions is called the *Pareto frontier*. More specifically, a language is (Pareto) optimal if there is no other language that has both lower complexity and higher informativeness. Computational modeling of cross-linguistic semantic data has demonstrated that natural languages optimize the simplicity/informativeness trade-off (Kemp and Regier 2012, Steinert-Threlkeld 2019, Denić et al. 2021, Uegaki 2020, Xu et al. 2020). Importantly, simplicity of a a language is assumed to be a function of the semantic representations underlying its expressions: these studies thus stipulate underlying semantic representations of languages' expressions, and analyze the optimality of the simplicity/informativeness trade-off under those stipulations. In the present work, we reverse the direction of the analysis: we assume that languages trade optimally simplicity and informativeness, which allows us to infer the semantic representations underlying their expressions. Concretely, the aforementioned 320 hypotheses will be evaluated as follows: as we assume that natural languages should be optimal solutions to the problem, we will have evidence against a hypothesis if under that hypothesis natural languages are not at the Pareto frontier. This approach connects to recent work by Zaslavsky et al. (2021), who show that two different hypotheses about cognitive biases involved in personal pronoun systems lead to different trade-off results.

Complexity of numeral systems We take the complexity of a combination of primitives to be the sum of weights of primitives and operations involved in the combination. For instance, if $w_1 = 1$ and $w_+ = 2$, the complexity of '1+1' is 4. We assume that the semantic representation underlying a numeral is the lowest complexity combination of primitives compatible with its denotation. The complexity of a language is defined as the sum of complexities of semantic representations underlying its numerals.

Informativeness of numeral systems The informativeness of a language $I(L)$ is defined in (6) (cf. Skyrms 2010, Steinert-Threlkeld 2019, Denić et al. 2021). It corresponds to the probability that the

Numeral systems	Languages
1, 2, 3	Bare, !Xóõ
1, 2, 3, more than 3	Achagua, Araona, Hixkaryana, Krenak, Mangarrayi, Martuthunira, Pitjantjatjara
1, 2, 3, 4, more than 4	Awa Pit, Kayardild
1, 2, 3, 4, 5, more than 5	Barasano, Imonda, Rama, Yidini
1, 2, 3, 4, 5, 6, 7, 8, 9, 10	Hup, Waskia, Wichi

Table 1: 18 exact restricted languages per their numeral systems inventory (Xu et al. 2020, Comrie 2013). For instance, Krenak has terms for 1,2,3, and a term that can be used for any number greater than 3.

communication will be successful given the speaker's probability to use an expression e to communicate a number n, $P_S(e|n)$, as in (4), the listener's probability to guess n upon hearing e, $P_L(n|e)$ as in (5), and the need probability to communicate about different numbers $P(n)$, which we assume to be approximated by a power law distribution as in (7), following Dehaene and Mehler (1992), Piantadosi (2016).

Natural languages Our natural language sample consists of the 18 exact restricted languages from Xu et al. (2020). They can be divided into 5 classes (Table 1) in terms of their numeral systems inventory denoting numbers 1–10. Informally, these are languages for which it's not the case that for every natural number they have a numeral non-ambiguously denoting it.

Hypothetical languages We generate all numeral systems ($N = 1534$) with (i) numerals denoting one of the numbers 1-10, and/or (ii) an expression meaning 'more than n', for some number n. We assume that expressions within a single language don't overlap in their denotations.

Results For each of the the 320 hypotheses: (i) we compute the complexity and informativeness of natural and hypothetical languages; (ii) we find the set of optimal languages (the Pareto frontier); (iii) we compute the average distance $\overline{D}$ of natural languages from the Pareto frontier. If languages are optimal solutions to the trade-off problem, we can discard all hypotheses for which $\overline{D} \neq 0$. We find that there are 10 out of 320 hypotheses for which $\overline{D} = 0$. These 10 hypotheses can be compressed into 2 families of hypotheses: (i) PRIM = $\{1\}$ and $w_s > w_>$; or (ii) PRIM = $\{1,2\}$ and $w_{\text{PRIM}}, w_+, w_s > w_>$. Interestingly, no hypothesis where PRIM contains $[1, \ldots, n]$, with $n \in \{3, \ldots, 10\}$, results in natural languages being Pareto optimal.

Discussion (i) The 310 hypotheses not resulting in the optimal simplicity/informativeness trade-off can only be discarded under the assumption that natural languages are optimal solutions to the trade-off problem. This assumption may be too strong: natural languages may be very good solutions, but not necessarily optimal. If this is the case, our approach cannot provide categorical evidence against certain hypotheses, but can nonetheless be used to evaluate their plausibility: if natural languages are far from the Pareto frontier under a specific hypothesis, that makes the hypothesis unlikely to be true. (ii) It would be interesting to explore more fine-grained weight assignments, which may reveal that under specific assumptions, PRIM other than $\{1\}$ or $\{1,2\}$ can result in the optimal simplicity/informativeness trade-off. (iii) The approach we develop adds to existing approaches to inferring semantic representations (cf. *Introduction*), creating novel opportunities to compare and integrate findings from multiple approaches.

Conclusion We have developed a new methodology for studying semantic representations underlying a semantic domain. We applied the method to numerals; importantly, the method can be applied to other semantic domains for which cross-linguistic semantic data is available. We thus hope that it will be a valuable tool for studying semantic representations underlying truth-conditional meanings.

(4) $$P_S(e|n) = \frac{[\![e]\!](n)}{\sum_{e' \in L} [\![e']\!](n)}$$

(5) $$P_L(n|e) = \frac{[\![e]\!](n)}{\sum_{n' \in N} [\![e]\!](n')}$$

(6) $$I(L) = \sum_{n \in N} \sum_{e \in L} P(n) P_L(n|e) P_S(e|n)$$

(7) $$P(n) \propto n^{-2}$$

Selected references: Denić et al. (2021). Complexity/informativeness trade-off in the domain of indefinite pronouns. *SALT 2020* | Fodor (1975). The language of thought. *HUP* | Kemp & Regier. Kinship categories across languages reflect general communicative principles. *Science* | Xu et al. (2020). Numeral systems across languages support efficient communication. *Open Mind* | Zaslavsky et al. (2021). Lets talk (efficiently) about us: Person systems achieve near-optimal compression. *CogSci 2021*