CHAPTER 11

Interpretation of numerals under memory load by Vietnamese speakers

Andreas Haida, Tue Trinh and Chi Mai Luong
Hebrew University of Jerusalem, Israel / Leibniz-Centre General Linguistics (ZAS), Germany / Institute of Information Technology, Vietnam Academy of Science and Technology, Vietnam

Numerals show an ambiguity between a weak, 'at least' meaning and a strong, 'exactly' meaning. The Gricean approach takes the weak meaning to be basic and derives the strong meaning as implicature, thus assimilates numerals to other scalar items. The Fregean approach, in contrast, takes the strong meaning of numerals to be basic and derives the weak meaning via type shifting operations. This paper gives a brief summary of these two approaches, followed by a report on a dual-task experiment which is designed to test how Vietnamese speakers interpret numerals under different memory loads. The goal of this experiment is to replicate the results of Marty et al. (2013) which can be interpreted as supporting the Fregean approach. It turns out that this goal could not be achieved, and we give some speculations as to why it was not.

Keywords: numerals, implicatures, scales, memory, Vietnamese

1. Theoretical background

Natural language numerals such as *three* are ambiguous between a "weak" meaning ('at least three') and a "strong" meaning ('exactly three'). Example (1) illustrates the weak meaning: replacing *three* with *at least three* has no effect, while replacing *three* with *exactly three* makes the sentence incoherent. Example (2) illustrates the strong meaning: replacing *three* with *exactly three* has no effect, while replacing *three* with *at least three* makes the sentence incoherent.

(1) a. John has three children, possibly four.
b. John has at least three children, possibly four.
c. #John has exactly three children, possibly four.

https://doi.org/10.1075/slcs.211.11hai

(2) a. John has three children, not four.
b. #John has at least three children, not four.
c. John has exactly three children, not four.

One obvious way to account for this ambiguity is to appeal to homophony. It is a fact about language that different words may have the same pronunciation, and to know which of the same sounding words is being used we have to consider the linguistic context. For example, the fifth syllable in (3a) is most probably the word which means 'a financial establishment that invests money deposited by customers' ($bank_1$) while the fifth syllable in (3b) is most probably the word which means 'the land alongside a river or lake' ($bank_2$).

(3) a. I went to the bank to cash a check.
b. I went to the bank to fetch some water.

We could say that there are two lexical items in English, $three_S$ and $three_W$, which mean 'exactly three' and 'at least three,' respectively. The linguistic context would resolve the ambiguity subject to pragmatic preferences, one of which would be the preference for the strongest non-contradictory meaning. Thus, a plain sentence such as (4) would be parsed with $three_S$, but (1a) would be parsed with $three_W$ and (2a) parsed with $three_S$, as only these parses yield a non-contradictory meaning.

(4) John has three children.

This approach faces several challenges. First, we would have to say that not only there are $three_S$ and $three_W$, but there are also $four_S$ and $four_W$, $five_S$ and $five_W$, and so on ad infinitum. Second, we would have to say that such pairs happen to be extremely popular across speech communities in the world, as the paradigm in (1) and (2) can be replicated in many, perhaps all, languages. Third, people learn these pairs naturally and without explicit instructions, quite differently from how they learn the various meanings of, say, *bank* in the English. Fourth, the different meanings in the case of numerals are systematically related in a way which is very different from how the meanings of *bank* and other homophonous word pairs are related. These facts should suffice as arguments that homophony is not the right answer. But the observation which decisively puts the homophony approach to rest is perhaps the coherence of the following sequence.

(5) John has three children. Bill does too. In fact, Bill has four children.

The first sentence clearly implies that John has exactly three children. Thus, it should be parsed with $three_S$. The second sentence, however, cannot be parsed with $three_S$, since such parse would bring it into contradiction with the third sentence, in the same way that the sequence in (6) is a contradiction.

(6) #Bill has exactly three children. In fact, Bill has four children.

This means that the first two sentences of (5) would have to have the following logical forms. PF deletion of the VP in the second sentence is represented by strikethrough.

(7) [$_{TP}$ John T_{PRES} [$_{VP}$ have three$_{S}$ children]. [$_{TP}$ Bill T_{PRES} [$_{VP}$ ~~have three$_{W}$ children~~]] too.

But this analysis is incompatible with a condition on VP-ellipsis, Parallelism, which requires that the elided VP be lexically identical with the antecedent VP (Fox 2003). Parallelism explains why the sequence in (8a) cannot mean that John went to the money bank and Bill went to the river bank: it cannot be parsed as in (8b), where the elided VP is not lexically identical to the antecedent VP.

(8) a. John went to the bank. Bill did too.
b. [$_{TP}$ John T_{PAST} [$_{VP}$ go to the bank$_{1}$]]. [$_{TP}$ Bill T_{PAST} [$_{VP}$ ~~go to the bank$_{2}$~~]] too.

Given Parallelism, the analysis in (7) is ruled out, which means the meanings 'at least three' and 'exactly three' cannot be derived from two different lexical items.

This strong/weak contrast in meaning exhibited by numerals turns out to be much more general. Quantifiers such as *some* and connectives such as *or*, for example, show the same variation.

(9) a. John did some of the homework but did not do all of them.
b. John did some of the homework. In fact, he might have done all of them.

(10) a. John talked to Mary or Sue but not both.
b. John talked to Mary or Sue. In fact, he might have talked to both.

Example (9) is evidence that *some* has a strong reading where it means 'some but not all' and a weak reading where it means 'some or all.' Similarly, (10) shows that *or* has an strong, "exclusive," reading where the disjunction is only true when exactly one disjunct is true, and a weak, "inclusive," reading where the disjunction is only false when both disjuncts are false. And as we can see from (11), the arguments against an explanation in terms of homophony for numerals can be replicated for these items also.

(11) a. John did some of the homework. Bill did too. In fact, Bill did all of the homework.
b. John talked to Mary or Sue. Bill did too. In fact, Bill talked to both of them.

Paul Grice, in his seminal Harvard lectures, sketched a path to understanding this phenomenon. The idea is to take the weak meaning to be basic and derive the strong meaning as "implicatures," i.e. inferences drawn on the basis of the literal meaning plus reasoning on the speaker's belief. The steps in this reasoning, Grice proposed, is justified by universal principles of rational communication, which he called "maxims of conversation." Together these maxims say that speakers, by default, assert the most informative proposition among those which are relevant and which they believe to be true. In a context where the question under discussion is how many children John has, (12a) and (12b) would both be relevant. Assuming the literal meaning of numerals to be the weak, 'at least' meaning, (12b) is more informative than (12a). This means, given the maxims of conversation, that a speaker who asserts (12a) does not believe that (12b) is true and therefore, assuming she knows how many children John has, believes that John does not have four children, i.e. that he has exactly three children.

(12) a. John has three children.
b. John has four children.

Similarly, a speaker who asserts (13a), given that both (13a) and (13b) are relevant, will convey the belief that John did not do all of the homework, and a speaker who asserts (14a), in the same way, will convey the belief that John talked to only one of Mary and Sue, under the assumption that both (14a) and (14b) are relevant.

(13) a. John did some of the homework.
b. John did all of the homework.

(14) a. John talked to Mary or Sue.
b. John talked to Mary and Sue.

The "Gricean account" of this fact takes the weak meaning to be basic and derives the strong meaning by way of scalar implicature, thus assimilating the weak/strong ambiguity of numerals to the well-known weak/strong ambiguity of other items such as quantifiers or connectives. It thus has the virtue of generalization, having subsumed various empirical observations under one phenomenon. However, it should be noted that there are other observations which speak against treating numerals in the same way as quantifiers and connectives. Specifically, numerals seem to retain their strong meaning in downward environments where the weak meaning would actually lead to a stronger interpretation for the sentence as a whole, and where, as expected, quantifiers and connectives show their basic, not derived, meaning (Horn 1972; Breheny 2008). As illustration, consider (15a) and (15b). Most speakers would interpret *pets or children* in (15a) as 'pets or children or both,' but would find it quite natural to interpret *three* in (15b) as 'exactly two' (Breheny 2008).

(15) a. Everyone who has pets or children is happy.
b. Everyone who has three children is happy.

In other words, we would immediately judge (15a) as false if it turns out that someone who has both children and pets is unhappy, but would not consider a miserable person who has four children to be a counterexample to (15b). In addition, experimental studies have provided evidence for children's differential treatment of numerals *vs*. quantifiers and connectives (Noveck 2001; Musolino 2004). Specifically, children are less prone than adults to the strong interpretation of quantifiers and connectives, but equally prone as adults to the strong interpretation of numerals. These observations have led some researchers to propose that for numerals, the strong reading is actually basic, with the weak meaning derived by way of type shifting operations. This approach is sometimes called the "Fregean approach," a term used by Kennedy (2015) which takes numerals to be properties of predicates, a view espoused by Frege (1884). Here is a simplified version of Kennedy's analysis.

(16) Kennedy's semantics for numerals
a. $[[\text{three}]] = \{P \mid \max(\{n \mid \exists x: P(x) \,\&\, \#_P(x) = n\}) = 3\}$
where max(X) is the largest number in X and $\#_P(x)$ is the number of P-atoms in x
b. [[John read three novels]] = 1 iff $\{x \mid x \in [[\text{novels}]] \,\&$ John read $x\} \in$ [[three]]

In plain English, *three* is the property of predicates which are true of three, but no more than three, entities, and *John read three novels* is true *iff* John read three, but no more than three, individual novels. For the weak meaning of three, Kennedy resorts to the successive application of BE and ι, two type shifting operators which are proposed by Partee (1987) to be part of the inventory of semantic interpretation rules on natural language and which, together, have the effect of removing the maximality requirement. We do not have to go into the details of how this machinery works. Suffice it to say that the following holds.

(17) [[John read ι(BE(three)) novels]] = 1 *iff*
$\exists x: x \in [[\text{novels}]] \,\&$ John read $x \,\&\, \#_{\text{novel}}(x) = 3$, i.e. *iff*
there exists a plurality of three novels that John read

The existential quantification leads to a weak interpretation of the numeral: even if John read four novels, there will be a plurality of three novels that he read, which means the sentence is true when John read more than three novels.

2. Experiment

To summarize the last section, numerals are observed to be ambiguous between a weak, 'at least' meaning and a strong, 'exactly' meaning. The Gricean approach takes the weak meaning as basic and derive the strong meaning as an implicature, assimilating numerals to other logical terms such as quantifiers and sentential connectives. The Fregean approach, in contrast, takes the strong meaning to be basic and derive the weak meaning by way of applying type shifting operations, in effect claiming that numerals are fundamentally different from quantifiers and connectives.

This controversy constitutes the background of Marty et al. (2013) (henceforth MSC), which reports the results of an experiment where native speakers of French are asked to perform two tasks simultaneously: (i) memorize a sequence of two letters (low memory load) or four letters (high memory load), and (ii) form a truth-value judgement about (the French counterpart of) sentences such as (18a) or (18b) with respect to various depicted situations.

(18) a. Some dots are red.
b. Four dots are red.

It was found out that in the case of the existential quantifier, the strong meaning is assigned less often under high memory load than under low memory load, while in the case of numerals, the strong meaning is assigned more often under high memory load than under low memory load. This finding supports the Fregean account of numerals, at least for French, since it is more compatible with the claim that for the existential quantifier, the weak meaning is basic and the strong meaning is derived, while for the numeral, the strong meaning is basic and the weak meaning is derived.

In the summer of 2016, the three authors of this paper ran an experiment on Vietnamese speakers which aims to replicate Marty *et al.*'s results. Note that Vietnamese is typologically unrelated to French and exhibit several properties in the domain of numeral phrases which distinguish it from French. Two examples are (i) that combination of a numeral and a nominal must be mediated by a "classifier," as illustrated in (19a), and (ii) that numeral phrases can be interpreted as indefinites only post-verbally, as illustrated in (19b).

(19) a. *John đã cắn hai *(con) chó.*
John PAST bite two *(CL) dog
'John bit (the) two dogs.'
b. *Hai con chó đã cắn John.*
two CL dog PAST bite John
'The two dogs bit John' / *'two dogs bit John.'

An analysis of these and other facts takes numerals in Vietnamese to have the weak meaning as basic and derives their strong meaning as an implicature. These considerations motivate running MSC's experiment, or variants thereof, on Vietnamese. Replication of the French results would further validate the dual-task paradigm as a way to gain insights into the processing of ambiguous expressions. A failure to replicate the French results would prompt further research into the linguistic differences between French and Vietnamese and possibly research into the cultural differences between the corresponding linguistic communities.

We implemented a software to run a dual task experiment with Vietnamese instructions and materials. We wrote the software in Python 3 (using PyCharm Community Edition) and we made heavy use of the matplotlib plotting library. The software will be published under the MIT open source license.

We ran three variants of MCS's dual-task experiment.[1] Figures 1 and 2 below show two example screens with material of the truth-value judgment task that we used. The caption in Figure 1 reads "in the circle there are less than five red stars." This is a control condition aimed at selecting participants who can count properly and thus do the task competently. The caption in Figure 2 reads "in the circle there are five blue stars." This is the target condition aimed to test which reading the numeral receives: a "true" answer indicates the weak reading, while a "false" answer indicates a strong reading.[2]

In Experiment 1 (a pilot study with 12 participants), we replicated MCS's design and method as closely as possible (i.e., modulo the differences induced by the structural peculiarities of Vietnamese). Specifically, we instructed the participants in the same way as MCS: we emphasized the importance of the memory task and encouraged them to rely on their intuition in the linguistic rating task (using a slider scale between 0 and 100). Notably, like MCS we did not mention that the expressions of interest (numerals and scalar items) are ambiguous between a weak and a strong reading.

1. We thank the following people, among many others, for their help in running our experiments: Trịnh Minh Giáp, Amber Ha Nguyen, Diep Vo, Thu Hien Pham, Nguyễn Hải Đăng, Sung Nguyen, Ngân Ngọc Tố Ngân, Thai Nguyet, Hieu Nguyen, Quỳnh Trang, Med Zed, and Su Sikei.

2. A reviewer points out that in Figure 2, the blue stars inside the circle are divided into a group of two and a group of five, and expresses concern that this might be a confound. In fact, MCS addresses this issue. The worry is that participants might interpret the sentence with a domain restriction which interpret "there are five stars" as 'there are a subgroup of five stars.' However, MCS has shown that the worry is unwarranted, and that participants do not engage in such domain restriction in any meaningful way. We find MCS's argument to be convincing.

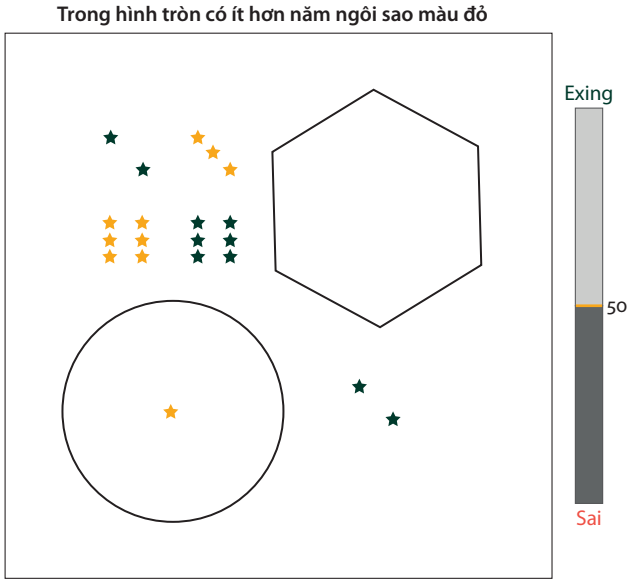

Figure 1. Example (1)

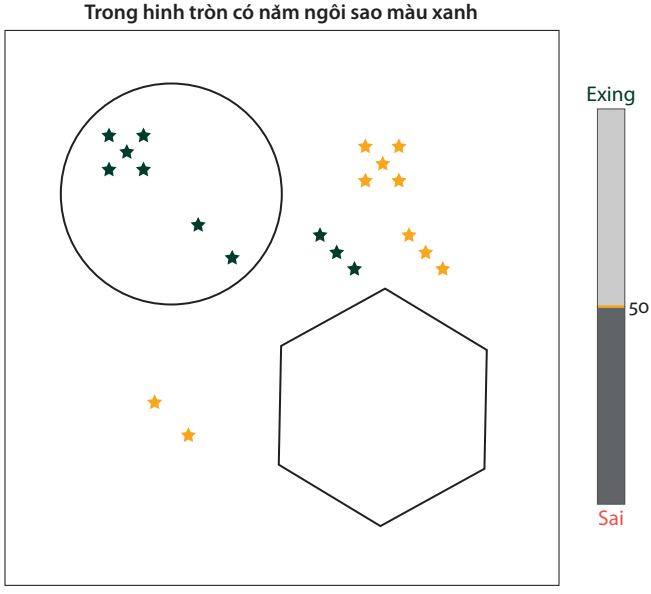

Figure 2. Example (2)

The results of the pilot experiment showed a significant effect of memory load for the numeral condition ($t = 1.98$, $p = 0.049$), strikingly in the *opposite* direction of MCS (see Figure 3). However, there was no effect for the Vietnamese counterparts of *some* and *or* (both ts < 0.8, n.s.) because of the ceiling effect (e.g., for *some*: mean = 93.67, SD = 7.65, skew = −1.42). That is, scalar items were overwhelmingly assigned the weak meaning.

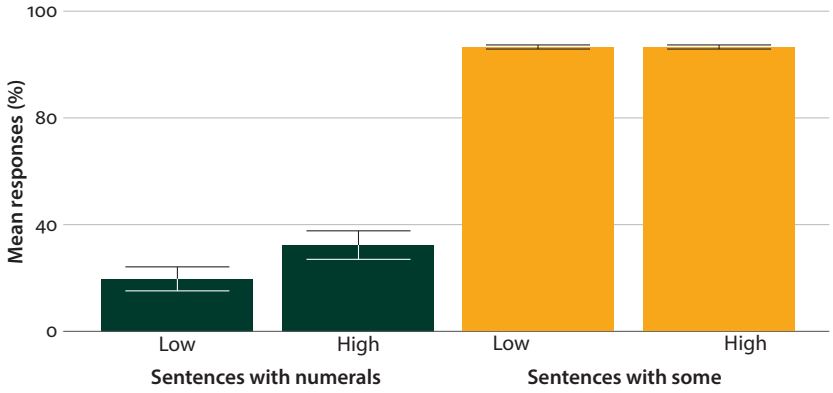

Figure 3. Results of Experiment 1

In Experiment 2 (32 participants), we changed the instructions to make the participants aware of the ambiguity of the expressions of interest with the ultimate goal to raise awareness of the weak reading of scalar items (see the instructions in the Appendix). This modification had an adverse effect on the results: no significant effect in any of the target conditions (all ts < 1.2, n.s.).

In Experiment 3 (31 participants), we modified the mode of presentation of the linguistic expressions in the rating task (auditory in addition to visual presentation). The rationale for this change was to reduce the processing load in the linguistic task to free up resources for scalar inference computation (again with the goal to make the strong reading of scalar items more readily available). This change led to a significant effect for the scalar item *or* ($t = 2.2$, $p = 0.03$), as can be seen from Figure 4. However, it had an adverse effect on the result with respect to numerals ($t = -0.16$, $p = 0.87$).

The results of our experiments differ from the result reported in MCS in various ways. Most significantly, none of our experiments evoked a significant effect in both of the crucial conditions: in experiment 1, the participants' behavior in the low memory condition and high memory condition did not differ significantly for sentences containing scalar items; in experiment 3, we did not find a difference for sentences containing numerals; experiment 2 did not evoke significant effects at

all. This means that our experiments fails to replicate the result of MCS. Thus, our results do not support the Fregean account of the ambiguity of numerals. However, neither do they support the Gricean account since we could not establish that constraining memory resources has an opposite effect on the processing of numerals than on the processing of scalar items. That is, our results are inconclusive.

What could be responsible for the different outcome of our experiment compared to that of MCS? A plausible hypothesis is that it is induced by the change in the experiment material. In our material, the critical phrase occurred in post-verbal position. In the French experiment, it appeared in pre-verbal postion. Therefore, in a first step to understand the difference in outcome we plan to perform a dual task experiment on French with French counterparts of our Vietnamese material. The result of this experiment will allow us to see if numerals and/or scalar items are processed differently in post-verbal position that in pre-verbal position.

Acknowledgements

This research was funded by the Vietnam Institute for Advanced Study in Mathematics.

References

Breheny, Richard. 2008. A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics* 25: 93–139. https://doi.org/10.1093/jos/ffm016

Fox, Danny. 2003. On logical form. In *Minimalist Syntax*, Randall Hendrick (ed.), 82–123. Oxford: Blackwell.

Frege, Gottlob. 1884. *Die Grundlagen der Arithmetik*. Breslau: Verlag von Wilhelm Koebner.

Horn, Laurence. 1972. On the Semantic Properties of the Logical Operators in English. PhD dissertation, UCLA.

Kennedy, Christopher. 2015. A "De-Fregean" semantics (and Neo-Gricean pragmatics) for modified and unmodified numerals. *Semantics and Pragmatics* 8: 1–44. https://doi.org/10.3765/sp.8.10

Marty, Paul, Chemla, Emmanuel & Spector, Benjamin. 2013. Interpreting numerals and scalar items under memory load. *Lingua* 133: 152–163. https://doi.org/10.1016/j.lingua.2013.03.006

Musolino, Julien. 2004. The semantics and acquisition of number words: Integrating linguistic and developmental perspectives. *Cognition* 93: 1–41. https://doi.org/10.1016/j.cognition.2003.10.002

Noveck, Ira A. 2001. When children are more logical than adults: experimental investigations of scalar implicature. *Cognition* 78: 165–188. https://doi.org/10.1016/S0010-0277(00)00114-1

Partee, Barbara. 1987. Noun phrase interpretation and type shifting principles. In Geroen Groenendijk, Dik de Jongh & Martin Stokhof (eds.), *Studies in discourse representation and the theory of generalized quantifiers*, 115–143. Dordrecht: Foris. https://doi.org/10.1002/9780470758335.ch15

Appendix

The following quote gives the English counterpart of the instructions of experiment 2. The instructions for experiment 3 reflected the changes in the experiment we describe in the last section.

> Thank you for your participation in our experiment. Each assignment in this experiment has three steps.
>
> (1) In the first step, you will see a sequence of letters appearing on the screen one after another. Your task is to remember these letters and their order of appearance.
> (2) In the second step, you will see a picture, containing a square, a circle, ten red stars, and ten blue stars. Above the picture is a sentence describing the picture. Your task is to judge how true or false this sentence is by choosing an appropriate position on the scale next to the picture. Use the up and down keys to make your choice. The top-most position indicates "completely true," and the bottom-most position indicates "completely false."
>
> When you are finished, press "enter" to submit your answer and continue to the next step.
>
> (3) In the third step, your task is to type in the letters in the reverse order of the sequence that you see in step (1). For example, if you see ABCD in step (1), you are to click DCBA now. After you have done this, press "enter." You will then see how well you did in step (3). When you are ready to go on to the next assignment, press "enter" again.
>
> (Please note that this is not a math test. We are not checking your mathematical knowledge. What we are interested in is your intuition about sentences in natural language. A wide-spread phenomenon in natural language is that sentences can be true or false to certain extent. For example, the sentence "February has 28 days" is true but not totally true, because there are leap years. The sentence "March has 28 days" is false but not totally false, because in a sense, every month has 28 days. Thus, feel free to use your intuition and choose intermediate points on the scale. Especially, do not try to answer based on what you have learned in math or logic classes. You should answer as quickly, and as spontaneously, as possible.)