

Some approximations: an experimental investigation

Stephanie Solt (ZAS), Jon Stevens (OSU) and Brandon Waldon (ZAS)

The problem. The sentence in (1) illustrates a curious use of *some* in which it combines with a numerical expression (the ‘*some* + numeral’ construction, or SN for short):

- (1) Some 20 cars were involved in the accident.

Authors including Sauerland & Stateva (2007) and Anderson (2014) align *some* on this use to approximators such as *about* and *approximately* (i.e. *some* 20 \approx *about* 20), an idea that may be formalized via the mechanism of scale granularity (per (2)) or pragmatic halos (per (3)):

(2) $\llbracket \text{some } twenty \rrbracket^{gran} = \text{coursest}(\text{gran})\llbracket twenty \rrbracket$ (Sauerland & Stateva 2007)

(3) $\llbracket \text{some } twenty \rrbracket^C = f(\llbracket twenty \rrbracket \cup \text{halo}_C(\llbracket twenty \rrbracket))$ (Anderson 2014)

However, such analyses fail to capture distributional restrictions that distinguish *some* from approximators such as *about*.

i) **Sum-based interpretation:** In contrast to true approximators, SNs are restricted to occurring with numerical expressions that can be interpreted as the sum or aggregation of some type of entity or unit of measure (e.g. pluralities, temporal/spatial extents, but not clock times):

- (4) a. The meeting lasted some 3 hours. / We drove some 30 miles. / *It’s some 3 o’clock.
b. It’s about/roughly/approximately 3 o’clock.

In support of this, a corpus analysis (COCA; Davies 2008-) found no examples of non-sum-based measures in SNs; for approximators these accounted for 3-5% of non-cardinality tokens.

ii) **Lack of degree interpretation:** SNs differ further from approximators in being infelicitous in mathematical statements and as answers to *how many?* questions (with no overt pronoun).

- (5) Seven times fourteen is about 100 / approximately 100 / roughly 100 / ??some 100.

- (6) Q: How many students passed the test? A: 50 / about 50 / ?some 50 / some 50 of them.

From this we conclude that SNs do not have denotations in the degree domain (contra (2),(3)), but rather interpretations that are more closely integrated with that of the following NP.

An even more central issue is that while some speakers attribute an approximative interpretation to SNs, others do not. This is reinforced by corpus examples in which SNs are used to convey a precise known value; (7b), for example, could not be paraphrased as ‘she bore him about 6 children’. This is unexpected on an analysis that treats *some* as an approximator.

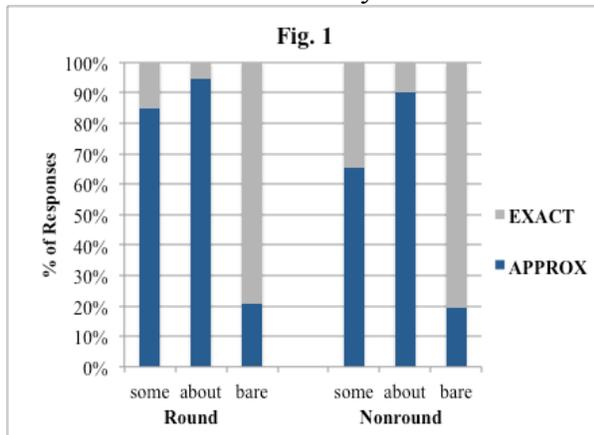
- (7) a. Of some 206 students who responded to the survey, 52% were female.
b. She bore him some 6 children, 3 of them boys.

Objectives. In order to develop a revised analysis of SNs that captures the patterns in (i) and (ii), it is necessary to understand the source of the variation in the presence of the approximative effect. We consider two hypotheses: **H1: Speaker Variation:** some speakers interpret *some* as an approximator, while others assign it a different meaning. **H2: Derived Approximation:** The approximative effect in examples such as (1) is not contributed by *some* itself, but rather derives from a possibility already available (if latent) for bare numerals. Krifka (2007) observes that round numbers allow approximate interpretations, while non-round numerals are interpreted precisely. If this is the source of the approximate interpretation of SNs, we predict *some*+round number to be interpreted approximately, but *some*+non-round to be interpreted precisely.

Experiment. To test the above hypotheses, native English speakers (n=75) were recruited via Amazon MTurk for an online interpretation study, in which they saw sentences such as (8) and were asked to indicate their interpretation of the numerical expression by providing a range:

- (8) The company added {about/some/∅} {50/47} new jobs in the first half of the year.
How many jobs did the company add in the first half of the year? Between ___ and ___.

A 3x2 design was employed, including 3 modifier conditions (*some*, *about*, bare) in 2 numerical conditions (round, non-round). Items were distributed across 5 lists: each respondent saw 1 item with a round number in all 3 modifier conditions, and a second item with a non-round number in all 3 modifier conditions. There were an additional 20 fillers, for a total of 26 items/participant. Presentation order was randomized for each participant. Data from 3 participants were excluded due to incomplete/inaccurate responses on filler items. The remaining responses were coded as EXACT (upper and lower values differ by at most 1 from stimulus value) or APPROXIMATE (upper and/or lower values differ by >1 from stimulus value).



As seen in Fig. 1, bare numerals elicited mostly EXACT responses and *about*+*n* primarily APPROXIMATE responses. *Some* constructions patterned distinctly from both, in the round condition eliciting mainly APPROXIMATE responses, but in the non-round condition exhibiting mixed behavior. A generalized logistic regression model found a significant difference between *some* and both bare ($z=7.8$, $p<0.001$) and *about* ($z=-4.2$, $p<0.001$). At the respondent level, the most common pattern (22/72) was to give APPROXIMATE responses to

‘about’ and ‘some’ trials in both round and non-round conditions, and EXACT responses to ‘bare’ trials in both conditions. But the second most common (16/72) was to treat *about* and *some*+round as APPROXIMATE but bare and *some*+non-round as EXACT. In total, roughly a third of participants distinguished *some* from *about* by giving *some* an EXACT interpretation in at least one condition where *about* received an APPROXIMATE interpretation.

Conclusions and preliminary analysis. The experimental results provide support for both of the above hypotheses. Some respondents do not differentiate *some* from the true approximator *about*. But others do (Speaker Variation), giving responses consistent with an analysis on which the approximative effect derives from a possibility inherent in the interpretation of bare round numbers themselves (Derived Approximation). We propose a two-part analysis: i) for some speakers, SNs are truth-conditionally equivalent to bare numerals, differing only in that they introduce a plural discourse referent formed via aggregation of atomic entities; in combination with round numbers they inherit the potentially approximate interpretation of the latter, an effect facilitated by the pragmatic principle of associating marked forms with marked meanings (Horn 1984); ii) for other speakers, this approximative meaning has been conventionalized into the semantics of *some* itself. In support of this, diachronic corpus data show that in earlier stages of English, SNs frequently (half of tokens) occurred with other markers of approximation (e.g. *some 15 or 20 men*), which we hypothesize to have encouraged semantic reanalysis.

REFERENCES. Anderson, C. 2014. Approximation of complex cardinals using *some*. *WECOL 2013*. Davies, M. 2008-. *The Corpus of Contemporary American English: 520 million words, 1990-present*. Horn, L. 1984. Towards a new taxonomy of pragmatic inference. In D. Schiffrin (ed.), *Meaning, Form, and Use in Context*. Krifka, M. 2007. Approximate interpretations of number words. In G. Bouma et al. (eds) *Cognitive foundations of interpretation*. Sauerland, U. & P. Stateva. 2007. Scalar vs. epistemic vagueness. *SALT 17*.