Between and Around

"Around" vs. "between". Consider a situation in which a speaker is asked how many people attended a party, but could not keep track of any precise number (see Williamson 1994 for a similar case involving a crowd). By assumption there is no number n such that the speaker can truthfully and justifiably assert :

(1)   There were exactly n people at the party.

For every value of n, either the sentence is going to be false, or it will be true without warrant. In both situations, uttering (1) would violate Grice's maxim of Quality (failure of true belief, or failure of adequate evidence). The speaker could, in principle, still abide by Grice's Quality maxim using precise intervals:

(2)   There were between n and m people.

In this paper we are interested in the reasons that typically make the specification of such intervals suboptimal compared to the use of a vague approximator (see Frazee and Beaver 2010), such as "around" in (3) (see Solt 2014; Krifka 2007):

(3)   There were around k people.

From the speaker's perspective, there appear to be two main problems with reporting only precise intervals. The first problem is that the intervals for which the speaker might be sufficiently confident to include the true value (in order to respect Grice's Quality maxim) may be too large to be informative, this time in violation of Grice's maxim of Quantity. A second problem, connected to the sorites paradox, is that there may not be a last stable value for which the speaker can be sufficiently confident that it is the last one for which she is sufficiently confident (see Williamson 1994). Intuitively, we appear to use "around" and other vague approximators in order to efficiently inform the hearer about our best estimate of the actual number. While it is admitted that "around n" triggers the implicature that "exactly n" is not assertible by the speaker (Solt 2014), the competition with exact "between" intervals is yet to be discussed.

A weak semantics for "around".   According to Krifka, and Solt, "around n" semantically denotes an interval of the form $[n - \frac{u}{2}, n + \frac{u}{2}]$, where u specifies a granularity. On the present proposal, "around n" has an existential meaning: it applies to x if x belongs to some interval or other of the form $[n - ku, n + ku]$. This semantics is very weak, but not vacuous. When the underlying scale is ratio and $u = 1$, the largest such interval is $[0, 2n]$. If the semantics is so weak, however, how can we explain the difference between uttering "x is around 20" and "x is between 0 and 40" in terms of informativeness? More generally, how can we explain the difference between "around j" and "between n and m" when j is equidistant between m and n?

When no contextual information interferes, we take is that "between n and m" conveys that all values in the corresponding interval are equally probable. By contrast, "around k" conveys that k is the most probable value to the speaker. How could a hearer draw that inference? One way to derive if from our weak semantics is to assume that a "literal listener" (Lassiter and Goodman 2017) puts a uniform probability distribution on the set of centered intervals of the form $[n - ku, n + ku]$. Under those assumptions, it can be seen that the resulting probability distribution for each value is Laplacean and peaks on n. Hence, the more remote a value from n, the less probable it is. If "between m and n" also manipulates a uniform distribution, then all values in $[m, n]$ are simply equiprobable, yielding a flat distribution instead.

Further predictions come with this proposal: with n increasing and keeping $u = 1$, the same level of probability encompasses more distant values to n (a prediction generally in line with Ferson et al. 2015's recent data). Provided the actual value m is not too far from n, "around n" will thus make m more likely to be considered by the hearer than asserting that the value lies in a single precise interval, while still minimizing the risk of error. "Around", in a situation of uncertainty, arguably yields a better tradeoff between Quality and Quantity than does an exact statement involving "between". Refinements can be considered to the model, finally, in which the lower and upper bound on the set of intervals for "around" are free parameters, in order to accommodate inter-individual variability and vagueness proper.

References

Ferson, S., O'Rawe, J., Antonenko, A., Siegrist, J., Mickley, J., Luhmann, C. C., Sentz, K., and Finkel, A. M. (2015). Natural language of uncertainty: numeric hedge words. *International Journal of Approximate Reasoning*, 57:19–39.

Frazee, J. and Beaver, D. (2010). Vagueness is rational under uncertainty. In Aloni, M., Bastiaanse, H., de Jager, T., and Schulz, K., editors, *Logic, Language and Meaning. Lecture Notes in Computer Science*. Springer.

Krifka, M. (2007). Approximate interpretation of number words. In Bouma, G., Krämer, I., and Zwarts, J., editors, *Cognitive Foundations of Communication*, pages 111–126.

Lassiter, D. and Goodman, N. D. (2017). Adjectival vagueness in a Bayesian model of interpretation. *Synthese*, 194(10):3801–3836.

Solt, S. (2014). An alternative theory of imprecision. In *Semantics and Linguistic Theory*, volume 24, pages 514–533.

Williamson, T. (1994). *Vagueness*. Routledge, London.