About 40 or so people: How and why we approximate

STEPHANIE SOLT

In his seminal book *The Number Sense*, psychologist Stanislas Dehaene relates the following joke:

> At the museum of natural history, a visitor asks the curator, »How old is this dinosaur over here?« »Seventy million and thirty-seven years« is the answer. As the visitor marvels at the accuracy of the dating, the curator explains: »I've been working here for 37 years, you know, and when I arrived I was told that it was 70 million years old«!
> (Dehaene 1997, p. 108)

This anecdote might not win any prizes at a stand-up comedy night, but it nicely illustrates an interesting point. Namely, when we communicate numerical information with one another, we often do so in approximate or coarse-grained terms. Obviously, whoever originally told the curator that the dinosaur skeleton was 70 million years old meant to convey ›approximately 70 million years‹. As competent speakers, we know that this is a possibility made available by our language, and we are able to infer other speakers' intended meaning when they make use of it.

In this brief contribution, we will survey some of the means natural languages have to convey approximate quantity and degree, and then delve into the deeper question of why this is often the preferred option.

Suppose we were asked how many people attended a particular lecture. The examples below illustrate some of the myriad ways a speaker of English might answer this question approximately. In the first example we see precision regulators or ›approximators‹ (Sauerland & Stateva 2007), words that serve to specify the degree of precision or imprecision at which a numerical expression should be interpreted. The second features disjunctive constructions, and the third a variety of nominal constructions.

(1) There were about / around / roughly / approximately 50 people there.
(2) There were 50 or 60 / 50 or so / more or less 50 people there.
(3) There were in the neighborhood of / on the order of / something like 50 people there.

English is by no means alone in having a rich vocabulary in this area: German has words such as *ungefähr*, *zirka*, and *rund*, Spanish has *alrededor de*, Mandarin *zuoyou*, and Japanese *gurai*, *yaku* and *oyoso*, among others. As an aside, the reader might note that many expressions of approximation derive from spatial terms: apparently we are able to think about the ›space‹ of numerical values via the metaphor of physical space.

Other languages have further possibilities. A nice example is Russian approximative inversion (Khrizman & Rothstein 2015). In Russian, a number word typically precedes the noun it quantifiers over, as in the first example below; but it may also follow the noun, as in the second example, with the effect that the number is understood as approximate:

(4) ona napisala desjat' knig
she wrote ten books
›She wrote ten books.‹
(5) ona napisala knig desjat'
she wrote books ten
›She wrote about ten books.‹

Furthermore, approximation can be conveyed even without any overt marker, simply by using a round number (Krifka 2007). For example, the utterance of the first sentence below would convey that exactly 99 people were present (if it turned out that the actual number was 98, we would conclude that the speaker was incorrect or lying). The second sentence, by contrast, allows or even prefers an approximate interpretation, similarly to what would obtain with an overt approximator such as *about*.

(6) There were ninety-nine people in the room.
(7) There were one hundred people in the room.

Road sign near Zurich airport

Our original dinosaur example makes the same point on a larger scale.

The wealth of means that languages have to express approximate number is itself evidence for the important role of approximation in communication and, potentially, cognition. But this point can even be made more directly: Dehaene and Mehler (1992) present data from numerous languages showing that round numbers are used much more frequently than non-round numbers of similar magnitude, a pattern they attribute to their use in conveying approximate quantities. Thus on a broad scale, approximation seems to be the more common option. In fact, in many contexts the use of a non-round number strikes us as comically over-precise, as in Krifka's (2007) example of a road sign near the Zurich airport which alerts drivers to the presence of a stop sign 103 meters ahead. In the remainder of this essay, we will consider possible explanations for this apparent preference for approximation.

One obvious reason to be approximate relates to **speaker ignorance**: speakers may choose an imprecise numerical expression because they lack precise knowledge of the facts, or because they are sufficiently uncertain of the facts to be willing to commit themselves to a precise description. I might for example tell you that the Palast der Republik was built ›in the mid 1970's‹ when I have no idea of the exact year, but also in the case that I think it was 1976, but I am not quite sure. In the latter situation, the choice to be approximate serves to avoid a violation of Grice's (1975) Maxim of Quality, which directs the speaker to try to make his conversational contribution one that is true.

Certainly, the speaker's knowledge state plays a role in many cases of numerical language use. To return to our original example, even the very latest techniques do not allow dinosaur skeletons to be dated with a precision of ±1 year. But there is also ample evidence that this cannot be the full story. The issue is that speakers often communicate approximately even when they do have precise knowledge. As an illustration, Dubois (1987) reports that scientists presenting their research to their peers frequently make use of approximations such as *about 40 normal subjects*, *about 45 millimeters of mercury*, *something like a six-fold increase* and *a bit greater than 1,1 millimeter*. Clearly, such choices of expression cannot be explained in terms of ignorance. A similar point is made by examples such as the following, seen frequently in the reporting of survey results and other sorts of statistical data.

(8) Roughly half (51%) of social media users have posted political content to their Facebook walls.
(9) About a third (34%) of online payments are now made on a mobile device.
(10) Four in five adults (79%) regard internet access as their fundamental right.

Here, the writer has the relevant precise data and furthermore chooses to report it, but uses an approximate expression as well. Apparently, reporting values or proportions in approximate or coarse-grained terms serves some communicative purpose that is not met by the communication of exact percentages.

As a second factor to consider, we might hypothesize that the preference for approximation is driven by a pressure towards the **reduction of speaker effort**. In other words, this would be to say that we tend to speak approximately simply because it is easier for us as speakers. Suppose I wish to tell you when a particular event took place. To say that it occurred ›a few weeks ago‹ is shorter than saying that it happened ›three weeks, two days and nine hours ago‹. A variety of theoretical frameworks concur that shorter expressions are preferred over longer ones on the grounds of speaker economy. Beyond this, the choice of the first approximate expression plausibly places less burden on the speaker's memory and reasoning capabilities (here, the ability to calculate when exactly the event in question occurred).

It would be difficult to argue that considerations of speaker economy don't play some role in this domain, given their influence in other aspects of language use. But again, this cannot be the full explanation, because speakers sometimes choose approximation even when this requires additional effort on their part. This is already evident from examples such as (8)–(10), where the writer has chosen to do something extra beyond reporting a precise value. This point is reinforced by a series of studies carried out by van der Henst et al. (2002), which demonstrate that in answering the question *What time is it?*, speakers often give answers rounded to the nearest 5-minute mark. The simple but clever methodology used in their studies was to approach individuals in public places such as train stations, and ask them the time. If respondents did not round off at all, we would predict on purely statistical grounds that 20% of their answers would be ›round‹, i.e. multiples of 5. In fact, round answers were much more frequent than this. Furthermore, not only was rounding frequent among analogue watch wearers (98% of all responses) for whom giving a rounded value is arguably easier than attempting to be precise, it was also observed though at lesser levels among wearers of digital watches (66% of responses), who must perform an extra step to

convert the digital read-out of their watch to a rounded and thus approximate temporal expression.

If speakers do not necessarily approximate out of ignorance or to save effort, perhaps they do so to create a more positive impression of themselves. In the language of the Third Wave of variationist sociolinguistics (Eckert 2012), this amounts to saying that speakers exploit the **social meaning** of an approximate expression, that is, the set of properties and stances of the speaker that it indexes.

Here, though, a variety of evidence points towards it being precise rather than imprecise forms that are associated with better perceptions of and outcomes for the speaker. One domain where this is documented is in consumer advertising. Xie & Konrod (2013) demonstrate that describing a car as offering a 10.2% reduction in carbon emission signals a higher level of competence on the part of the advertiser than when the same car is described less precisely as offering a 10% reduction, though the effect is found primarily among subjects with low scepticism towards advertising.

A similar pattern has been found in research into the psychology of negotiating. As shown by Mason et al. (2013) and others, negotiations are most commonly started with a round number first offer. But non-round first offers are stronger anchors, leading on average to more conciliatory – that is, closer – counteroffers. Furthermore, the evidence is that this effect is mediated by perceptions of the individual making the offer: a party in a negotiation who makes a precise (non-round) offer is perceived as being better informed and prepared than one who offers a round number to start.

Finally, even the granularity at which a numerical value is expressed can impact perceptions of the speaker and thus the credibility of his or her claim. Zhang & Schwarz (2012) for example found that predictions stated using finer grained temporal units are perceived to be more accurate than those stated using coarser grained units. In one example from their experiments, when the expected duration of a construction project was described as *1 year*, subjects inferred a confidence interval of 140 days; but this decreased to 105 days when the duration was described as *12 months*, and to 84 days when it was described as *52 weeks*.

Yet throughout this and other work, there are also indications that more precise utterances are not always associated with more positive perceptions of the speaker. Both Zhang & Schwarz (2012) and Loschelder et al. (2016) find that precise estimates and offers can actually have a negative effect when the degree of precision exceeds the speaker's assumed competence level, and when the speaker is perceived as untrustworthy. In more linguistically oriented work, Beltrama (2017) finds that the choice of a precise numerical expression is

associated with positive speaker attributes such as ›intelligent‹, ›articulate‹ and ›reliable‹ but also negative attributes such as ›pedantic‹, ›uptight‹ and ›annoying‹. Preliminary research at the ZAS in the context of the ETAPS project (DAAD 57316845; collaboration with Dr. Heather Burnett, CNRS) suggests that such effects are dependent on the situation of utterance: whereas in a formal context, higher precision might convey positive speaker attributes, in an informal context such as a blog post, it is rounded or imprecise forms that are associated with positive dimensions such friendliness, while non-round forms may convey that the speaker is arrogant or a show-off. Thus there is reason to believe that speakers sometimes do choose an approximate numerical form with the goal of conveying a desirable impression of themselves.

A final hypothesis that we will explore is that the tendency for speakers to round or approximate is driven by hearer-oriented considerations, namely a desire to **reduce processing effort** on the part of the recipient of the information. Van der Henst et al. (2002), whose clock-time experiment was discussed above, suggest this as an explanation for their findings. Approaching the problem from the perspective of the theory of communication known as Relevance Theory (Sperber & Wilson 1986), they propose that a rounded answer to the question *what time is it?* is optimally relevant, in that it produces the same cognitive benefit as a precise one, but with a lower processing cost. They support the view that rounding represents a hearer-oriented strategy by means of further experiments which demonstrate that speakers are sensitive to the level of precision relevant to the hearer: while speakers commonly round off to the nearest 5-minute mark when reporting the time, they do so significantly less frequently when the requestor is perceived to require a more precise answer, for example when he is setting his watch, or claims to have a meeting a short time later

Solt et al. (2017) provide more direct evidence that rounded values – specifically clock times – are indeed easier to process than non-round ones. In two experiments designed to simulate ordinary, everyday tasks that speakers might need to carry out with numerical information, they show that round clock times (multiples of 5) are recalled and manipulated more quickly and accurately than their non-round counterparts. For example, subjects perform better in determining what is 45 minutes before 8:15 than in calculating 39 minutes before 8:21. Such results suggest that speakers are in fact helping hearers when they make the (small) extra effort to provide numerical information in a rounded form.

Looking more broadly, there is reason to think that it is particularly natural for approximation to be easier for us than precision. A large body of research in cognitive psychology has demonstrated that in addition to the ability to represent precise number, humans also

possess a second system known as the ›approximate number system‹ (ANS), which allows us to represent and reason with quantities approximately (Dehaene 1997; Gallistel and Gelman 2000; Feigenson et al. 2004). This system is essentially analog in nature, with (approximate) number encoded as patterns of activation on the equivalent of a ›mental number line‹. These representations can serve as the basis for comparison of quantities, and for simple approximate arithmetic such as addition and subtraction. The ANS is developmentally and evolutionarily more basic than the ability to represent precise number. It is operational not just in adults who possess a system of number words, but also preverbal infants, as well as members of societies without complex number systems. Non-human animals, too, show approximate numerical abilities. Furthermore, the ANS is engaged even in tasks involving precise quantities presented symbolically. For example, subjects are faster in judging that 9 is bigger than 3 than in judging that 6 is bigger than 5 – an example of the ›distance effect‹ that is the hallmark of the ANS. Thus plausibly, approximate numerical expressions are easy because they are most readily mapped to the representations generated by the ANS, our most basic, default way of conceptualizing quantities.

To conclude, just as languages offer many ways of speaking approximately, so too are there many reasons speakers may choose to do so: out of ignorance, to save effort or ease the hearer's processing, and in some situations to convey a particular image or stance. In many cases we might think that *Be Precise!* is a motto to live by – but in others we would be equally justifying in heeding the advice *Be Approximate!*

REFERENCES

Beltrama, Andrea. 2017. Iconicity and (im)precision: How pragmatic variation can convey social qualities. Workshop ›Social Meaning‹, Leibniz-Zentrum Allgemeine Sprachwissenschaft, Berlin, 27 August 2017.

Buckley, Paul B. and Clifford B. Gillman. 1974. Comparisons of digits and dot patterns. *Journal of Experimental Psychology* 103, 1131–1136.

Dehaene, Stanislas. 1997. The number sense: How the mind creates mathematics. Oxford: Oxford University Press.

Dehaene, Stanislas and Jacques Mehler. 1992. Cross-linguistic regularities in the frequency of number words. *Cognition* 43(1), 1–29.

Dubois, Betty Lou. 1987. »Something on the order of around forty to forty-four«: Imprecise numerical expressions in biomedical slide talks. *Language in Society* 16, 527–541.

Eckert, Penelope. 2012. Three waves of variation study: The emergence of meaning in the study of socioliguistic variation. *Annual Review of Anthropology* 41, 87–100.

Feigenson, Lisa, Stanislas Dehaene, and Elizabeth Spelke. 2004. Core systems of number. *Trends in Cognitive Science* 8, 307–314.

Gallistel, C. R., and Rochel Gelman. 2000. Non-verbal numerical cognition: From reals to integers. *Trends in Cognitive Science* 4, 59–65.

Grice, Paul. 1975. »Logic and conversation«. In: Peter Cole and Jerry L. Morgan (eds.), *Syntax and semantics. 3: Speech acts*, pp. 41–58. New York: Academic Press. pp. 41–58.

Khrizman, Keren and Susan Rothstein, 2015. Russian approximative inversion as a measure construction. In: Gerhild Zybatow, Petr Biskup, Marcell Guhl, Claudia Hurtig, Olav Mueller-Reichau, and Maria Yastrebova (eds.), *Slavic Grammar from a Formal Perspective. The 10th Anniversary FDSL Conference, Leipzig 2013*, pp. 259–272. Frankfurt am Main: Peter Lang.

Krifka, Manfred. 2007. Approximate interpretation of number words: a case for strategic communication. In Gerlof Bouma, Irene Krämer and Joost Zwarts (eds.), *Cognitive Foundations of Interpretation*, 111–126. Amsterdam: Koninklijke Nederlandse Akademie van Wetenschapen.

Loschelder, David D., Malte Friese, Michael Schaerer, and Adam D. Galinsky. 2016. The Too-Much-Precision Effect: When and Why Precise Anchors Backfire With Experts. *Psychological Science*, published online October 27, 2016, pp. 1562–1572.

Dr. Stephanie Solt

Mason, Malia F., Alice J. Lee, Elizabeth A. Wiley, and Daniel R. Ames. 2013. Precise offers are potent anchors: Conciliatory counteroffers and attributions of knowledge in negotiations. *Journal of Experimental Social Psychology* 49(4), 759–763.

Sauerland, Uli & Penka Stateva. 2007. Scalar vs. epistemic vagueness. In Masayuki Gibson and Tova Friedman (eds.), *Proceedings of Semantics and Linguistic Theory (SALT) 17*, 228–245. Cornell U: CLC Publications.

Sperber, Dan and Deirdre Wilson. 1986. *Relevance: Communication and Cognition*. Oxford: Blackwell.

Solt, Stephanie, Chris Cummins & Marijan Palmović (2017). The preference for approximation. *International Review of Pragmatics* 9(2), 248–268.

Van der Henst, Jean Baptist, Laure Carles and Dan Sperber. 2002. Truthfulness and relevance in telling the time. *Mind and Language* 17, 457–466.

Xie, Guang-Xin and Ann Kronrod. 2013. IS THE DEVIL IN THE DETAILS? The Signaling Effect of Numerical Precision in Environmental Advertising Claims. *Journal of Advertising* 41(4), 103–117.

Zhang, Y. Charles and Norbert Schwarz. 2012. How and Why 1 Year Differs from 365 Days: A Conversational Logic Analysis of Inferences from the Granularity of Quantitative Expressions. *Journal of Consumer Research* 39(2), 248–259.

DR. STEPHANIE SOLT

is a member of FBIV »Semantics and Pragmatics«, where she is principal investigator on the DFG-funded project Degree Attenuators: The semantics of attenuating polarity items in the degree domain. Her research focuses on the linguistic expression of quantity, amount and degree, and on issues of vagueness and imprecision in natural language. Solt holds a B.A. in Mathematics from Bryn Mawr College and a Ph.D. in Linguistics from the City University of New York. Prior to the ZAS, she has held positions at the University of Amsterdam and the State University of New York, New Paltz.