The Preference for Approximation

Stephanie Solt\textsuperscript{a}
Chris Cummins\textsuperscript{b,1}
Marijan Palmović\textsuperscript{c}

\textsuperscript{a}Zentrum für Allgemeine Sprachwissenschaft
Schützenstrasse 18
10117 Berlin
Germany
stephanie.solt@gmail.com
+49 30 20192 570
Corresponding Author

\textsuperscript{b}SFB 673 - Alignment in Communication
Bielefeld University
33615 Bielefeld
Germany
c.r.cummins@gmail.com

\textsuperscript{c}University of Zagreb
Laboratory for Psycholinguistic Research
Zvonimirova 8
Zagreb HR10000
Croatia
palmovic@erf.hr

\textsuperscript{1}Present address:
School of Philosophy, Psychology and Language Sciences
The University of Edinburgh
Dugald Stewart Building
3 Charles Street
Edinburgh, EH8 9AD
UK

Word Count: 2991
Abstract. There is a widespread tendency for numerical information to be communicated in a rounded or approximate rather than precise manner, a pattern argued to result from a desire to minimize processing effort for the hearer. We report two experiments demonstrating a processing advantage for round vs. non-round clock times, and providing evidence for the role of domain-specific scale granularity.

Key Words: Imprecision, granularity, time, language processing, language use

1. Introduction

In communication, precision is typically assumed to be preferable to imprecision. However, when dealing with numerical information, a variety of evidence points instead to a speaker and hearer preference for approximation over precision.

Round numbers such as 100 and 40 tend to be interpreted approximately, while non-round numbers such as 99 and 43 are interpreted precisely (Krifka 2002, 2007). Dehaene and Mehler (1992) present data showing that round numbers are used much more frequently than non-round numbers cross-linguistically, which they attribute to this interpretative preference. Moreover, 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 alerting drivers to the need to stop 103 meters ahead.

Speakers give round answers even when they have more precise knowledge. Van der Henst, Carles and Sperber (2002) demonstrate experimentally that respondents to “What time is it?” tend to give answers rounded to the closest 5-minute mark (e.g. 3:10 instead of 3:08). This applies even to individuals using digital watches, for whom rounding is presumed to be more effortful than reading out the precise time (see also Gibbs & Bryant 2008).

Some uses of approximation might be attributed to the speaker’s uncertainty as to the accuracy or reliability of their information, but this is not always the case. In written materials such as the reporting of survey results, approximate expressions of number or proportion often occur alongside an exact value, as in (1) (cf. Williams & Power 2009).

(1) Six in ten Americans (59%) read the Bible at least occasionally.

Apparently, reporting values in approximate or coarse-grained terms serves some communicative purpose that is not met by the use of exact figures. This pattern arises not just in informal discourse but even in how scientists communicate with their peers (Dubois 1987).

Why are speakers willing to exert extra effort to give approximate, rather than precise, information? An intuitive answer is that rounding somehow makes comprehension easier for the hearer. Van der Henst et al. (2002) frame such an explanation within Relevance Theory (Sperber & Wilson 1987, 1995), which affirms that speakers aim to produce utterances that are maximally relevant from the hearer’s perspective. Relevance here is a function of cognitive effect and processing cost: maximizing relevance involves minimizing the cognitive cost of processing, and
maximizing the cognitive effect that processing the utterance achieves. For telling the time, Van der Henst et al. propose that a rounded answer produces the same cognitive benefit as a precise one, but with a lower processing load. “Suppose you have an appointment at 3:30 p.m. and it is 3:08. Being told ‘It is 3:10’ is likely to be optimally relevant: the two-minute departure from the exact time is unlikely to have any consequences, and the rounded answer is easier to process.” (p. 464; emphasis added). Similarly, Krifka (2007) suggests that coarser-grained representation of quantitative information might be “cognitively less costly.”

However, neither account justifies the underlying assumption of a cognitive advantage for roundness or coarse-grainedness, and little prior research bears on this point. The most relevant work is perhaps Mason, Healy and Marmie (1996), who found that subjects better recalled round numbers (e.g. 11,000) than non-round numbers (11,365), even when tested only on the first two digits. This is consistent with the hypothesized processing advantage for roundness, but hardly conclusive.

A further question concerns what aspect of a linguistic form, or its meaning, causes it to be favored. One possibility is that the crucial property is roundness itself. Jansen and Pollmann (2001) operationally define roundness in terms of divisibility: specifically, whether a number is a single-digit multiple of a power of 10, or 2-, 2.5- or 5-times a power of 10. For example, 100 can be expressed in all of these ways (\(1 \times 10^2\); \(2 \times (5 \times 10^1)\); \(4 \times (2.5 \times 10^1)\); \(5 \times (2 \times 10^1)\)), and hence is maximally round. By contrast, 40 is less round, as it can be expressed in only three of these ways (\(4 \times 10^1\); \(8 \times (5 \times 10^0)\); \(2 \times (2 \times 10^1)\)), and 45 is even less round, being expressible in just one of these ways (\(9 \times (5 \times 10^0)\)). 43 has none of these divisibility properties, and is thus non-round.

Using round numbers of this type might plausibly lead to processing advantages for several reasons. As noted earlier, such numbers occur more frequently, and frequent expressions are well-known to impose lower processing loads (Oldfield & Wingfield 1965, Balota & Chumbley 1984, and many others). They are privileged in their representation in Arabic numerals (the roundest ending in 0). Their verbal forms tend to be shorter (e.g. ‘one hundred’ – 3 syllables vs. ‘one hundred and three’ – 5 syllables), which could confer advantages in some processing tasks (e.g. memory span; Baddeley, Thompson & Buchanan, 1975). Finally, Dehaene and Mehler (1992) suggest that round numbers might be directly mapped to extra-linguistic representations of quantity.

However, Krifka (2007) notes that rounder numbers are not necessarily shorter or simpler than their less round counterparts: forty-five, for example, is no shorter than forty-six, and one hundred is longer than ninety. Furthermore, roundness per se is not necessarily the driving factor behind imprecise interpretation. Although 45 is less round (in Jansen and Pollmann’s sense) than 40, 45 minutes seems more likely than 40 minutes to be used to convey an approximate meaning.

Consequently, Krifka proposes that the crucial property is not roundness itself but rather scale granularity. Measurements can be reported with respect to scales that differ in how coarse- or fine-grained they are, i.e. in the density of their scale points. For example, relative to a fine-grained scale of miles, the distance between Amsterdam and Milan might be reported as 514
miles, but relative to a coarser-grained scale based on 100-mile units it might be reported as 500 miles. Typical scale granularity levels use powers of ten (10-20-30-...; 100-200-300-...; etc.) or apply an operation of halving (5-10-15-...) or doubling (20-40-60-...) to such structures, making their scale points by definition ‘round’. But in certain domains other structures are observed, and in these cases the distinction between granularity and roundness can be appreciated. A prime example is the measurement of time: Krifka notes that possible granularity levels include hours, half-hours, quarter-hours, 5-minute increments, and minutes. The last three of these are represented in Fig 1. Relative to the 1-minute-granularity scale (a) in Fig. 1, the time displayed might be described as 2:41. But speaking more approximately, that same time might be described as 2:40 (5-minute granularity level, (b)) or even 2:45 (15-minute level (c)).

Thus an alternate hypothesis posits a processing advantage for values that occur on coarser-grained scales (e.g. 2:45) over those that occur only on finer-grained scales (e.g. 2:41). As with roundness, there are various reasons why this advantage might arise: such values are presumably more frequent (for example, events more commonly start at the quarter-hour than at times such as 2:41), and might be more easily related to non-linguistic representations of measurement (e.g., to marks on the analog clock face). Alternatively, one might argue that the more coarse-grained segmentation of continuous reality is itself cognitively preferable, making coarser-grained measurements inherently easier to process.

![Figure 1. Alternate granularity levels for time](image)

2. Experiments
We report two experiments designed to investigate the following questions:

1) Is there a processing advantage for ‘rounder’ numerical expressions?

2) If so, is it driven by:
   a. Numerical roundness in the sense of Jansen and Pollmann (2001)?
   b. Domain-specific scale granularity?
We focus on clock times, both because these have been the subject of prior research on rounding (Van der Henst et al. 2002, Gibbs & Bryant 2008), and because this domain enables us to tease apart numerical roundness and scale granularity.

2.1 Experiment 1
Experiment 1 was a short-term memory task, employing the Sternberg paradigm (Sternberg 1966).

2.1.1 Participants
34 native German speakers (mean age 26.6, 25 female), were recruited from the students body of Bielefeld University and Humboldt University Berlin, or by word of mouth in the Berlin area. Data from 4 further participants were excluded due to low accuracy or misunderstanding of the task. Participants were paid 5 euros for participation.

2.1.2 Materials
Our experimental stimuli consisted of sequences of 3, 4 or 5 clock times, described to participants as departure times for trains, followed by a probe time; the participants’ task was to decide whether the probe occurred in the sequence. For each sequence length, 10 items were constructed (5 correct/5 incorrect), each of which was implemented at three granularity levels, corresponding to scales based on units of 1, 5 and 15 minutes. For each granularity level, three specific minute values were selected as the basis for the stimuli:

- **Coarse** (15-minute granularity) - :15, :30, :45
- **Medium** (5-minute granularity) - :10, :25, :40
- **Fine** (1-minute granularity) - :21, :36, :51

The following shows a sample item in its 3 versions:

Coarse: 2:45 6:30 8:15  Probe: 2:30  
Medium: 2:10 6:25 8:40  Probe: 2:25  
Fine: 2:36 6:51 8:21  Probe: 2:51

90 trials were administered in total (3 sequence lengths x 10 items x 3 granularity levels).

2.1.3 Procedure
The task was administered on a PC running E-Prime. In each trial, participants saw the sequence of times (each shown for 2 seconds), followed by a 0.5-second pause, then the probe time. Participants pressed one of two keys to indicate whether or not the probe occurred in the sequence. The experiment comprised three blocks of different sequence length (3 item, then 4
item, then 5 item), with breaks between blocks. Trial order was randomized within blocks. There were 3 practice trials at the beginning of the experiment. Participants’ responses and reaction times were recorded.

2.1.4 Results
Trials with reaction times less than 500ms or greater than 10,000ms were removed (11 trials, 0.4%). Rates of incorrect responses, and the log-reaction time (logRT) for correct responses, are shown in Figures 2 and 3.

The effect of granularity on accuracy was assessed by fitting a generalized linear mixed-effects model (Baayen, Davidson & Bates, 2008), with granularity as a categorical fixed effect (levels: coarse, medium and fine; coarse coded as baseline) and sequence length as a numerical fixed effect; random effects for subject and item were included. Accuracy declined significantly with increasing sequence length (z = –3.024, p<0.01). Crucially, respondents were significantly less accurate for fine vs. coarse granularity (z = –3.555, p<0.001), but there was no significant difference between coarse and medium granularity. There was a significant interaction between fine granularity and sequence length (z = 3.206, p<0.01), with the difference between fine and coarse granularity level attenuated at sequence length 5.

The effect of granularity on LogRT for correct trials was assessed by fitting a linear mixed-effects model, obtaining p-values by the Markov Chain Monte Carlo (MCMC) method (Baayen et al. 2008). Fixed effects were granularity and sequence length (as above) and correct response type (yes/no); random effects were subject and item. Reaction times increased significantly with sequence length (p<0.001) and were longer for ‘no’ responses (p<0.001). Crucially, reaction times were longer for fine vs. coarse granularity (p<0.01), but no significant difference was observed between medium and coarse granularity. The model was not improved by adding interaction terms.
Figure 2. Experiment 1 – % Incorrect
2.1.5 Discussion
The results of Experiment 1 demonstrate an effect of roundness. ‘Round’ clock times such as 4:15 are recalled more accurately, and elicit faster responses, than ‘non-round’ times such as 4:21. This effect was attenuated for longer sequence lengths (as seen in the length-granularity interaction), suggesting that participants may have developed strategies during the experiment. However, the fact that the effect is observed most clearly in the first block of trials (sequence length 3) speaks to its robustness.

Our experiment included times at three distinct granularity levels, coarse (15-minute granularity), medium (5-minute) and fine (1-minute). The times at the coarse and medium levels were both round in the sense defined above; those at the fine level were non-round. The difference in accuracy and reaction time between coarse and fine (i.e. round vs. non-round), taken together with the absence of a difference between coarse and medium (both round), seems to indicate that the processing advantage is driven by roundness rather than granularity of representation.
However, the design of this experiment did not force respondents to process the stimuli as clock times. For example, 4:20 could have been stored and processed as the numeral 420. In this case, granularity levels specific to the domain of clock times would not be expected to play a role. We address this possibility in Experiment 2.

2.2 Experiment 2
Our second experiment used a novel “clock time addition/subtraction” paradigm, which required participants to manipulate stimuli as times rather than numbers.

2.2.1 Participants
Participants were 22 native German speakers (17 female, mean age 24.4), recruited from the student body of Humboldt University and/or by word of mouth in Berlin. Participants were paid 4 euros for participation.

2.2.2 Materials
Test items were clock time addition/subtraction problems, each consisting of a start time, an operation (plus or minus) and an increment time in minutes:

A) 3:45
   plus
   30
   -----  
   4:15

B) 6:48
   minus
   26
   -----  
   6:12

Each problem was followed by a candidate answer; the participants’ task was to decide whether it was correct (as in A) or incorrect (as in B).

720 test items were created, in which the following factors were varied:

i) Start time granularity level:
   - Coarse (15-minute): minute-place :00, :15, :30, :45
   - Medium (5-minute): minute-place :05, :20, :35, :50
   - Fine (1-minute): minute-place :03, :18, :33, :48
ii) Increment granularity level:
   - Coarse (15-minute): 30, 45
   - Medium (5-minute): 25, 40
   - Fine (1-minute): 26, 41

iii) Operation: plus or minus

Half of the items featured correct and half incorrect answers (+/− 5 minutes or +/- 10 minutes, such that participants could not answer correctly based on parity alone). Critically, for some items (such as A), the correct answer ‘spilled over’ to the next or previous hour compared to the start time, ensuring that participants had to process the stimuli as times rather than numbers.

2.2.3 Procedure
The task was administered on a PC running E-Prime. Participants read the instructions on the screen, and completed three practice trials. They then completed 3 blocks of 48 trials drawn randomly from the list of 720 test items. In each trial, participants saw a fixation cross for 2 seconds, followed by the start time for 1 second, and then the operation and increment time for 1.5 seconds, and then the end time. Participants pressed one of two keys to indicate whether the end time was correct or incorrect. Responses and reaction times were recorded.

2.2.4 Results
Before analysis, one outlier trial with reaction time <200 msec was removed.

Results are shown in Figures 4 and 5. For accuracy, a linear mixed-effects model was fitted with start granularity, increment granularity and spillover (yes/no) as fixed factors, and subject and item as random factors. Relative to coarse start granularity, respondents were significantly less accurate for medium (z=−2.19, p<0.05) and fine start granularity (z=−5.43, p<0.001). Similarly, relative to coarse increment granularity, respondents were significantly less accurate for medium and fine increment granularity (z=−2.84, p<0.01 and z=−3.14, p<0.01, respectively). Spillover decreased accuracy (z=−4.49, p<0.001). A significant interaction was found between spillover and medium increment granularity (z=2.40, p<0.05), reflecting a smaller effect of increment granularity in the spill trials. Adding further interaction terms did not improve the model.
Figure 4. Experiment 2 – % Incorrect by Granularity

For LogRT, a linear mixed-effects model was fitted, with p-values obtained via the MCMC method. Start granularity, increment granularity, spillover (yes/no) and operation (plus/minus) were included as fixed factors, and subject and item as random factors. Respondents were significantly slower in spillover trials (pMCMC<0.001), and for minus operations (pMCMC<0.01). More crucially, relative to coarse starting granularity, LogRT was significantly greater for medium and fine start granularity; similarly, relative to coarse increment granularity, LogRT was significantly greater for medium and fine increment granularity (all comparisons pMCMC<0.001). Finally, significant interactions were found between fine increment granularity and medium and fine start granularity (pMCMC< 0.05 and pMCMC<0.001, respectively), indicating that fine increment granularity resulted in smaller increases in LogRT at finer start granularities. Adding further terms did not significantly improve the model.
2.2.5 Discussion

As in Experiment 1, participants responded more accurately and faster for round vs. non-round times. However, here we also found a difference between coarse and medium start granularity (both round). Hence, scale granularity exerted additional influence beyond that of roundness (per our second research question).

Increment granularity also played a role: participants were more adept at ‘moving’ forwards and backwards along the time scale in ‘coarse-grained’ units. This effect weakened at finer start granularity levels, suggesting that participants who were forced to adopt a finer-grained representation incurred reduced costs in calculating fine-grained increments.

3. Conclusions

Our results support the hypothesis that rounding represents a hearer-oriented strategy aimed at reducing comprehension effort. ‘Rounder’ clock times are processed more accurately and quickly than non-round times. Thus in selecting such expressions, speakers aid their
interlocutors, by presenting information in a form that can be more easily recalled and manipulated.

We further found evidence that it is not only round numbers *per se* that are associated with greater ease of processing; a further advantage accrues to those values that occur on scales of coarse granularity, such as time expressions for quarter hours.

A question that arises from these findings is whether approximation or coarse granularity itself – as opposed to the numerical expressions corresponding to points on such scales – is advantaged in processing. This is not directly addressed by our research: both of our experiments created contexts in which time expressions are necessarily interpreted exactly, namely departure times for trains (Exp. 1) and mathematical operands (Exp. 2). Thus the comparative processing of precise vs. approximate meanings represents a subject for future research.

References


