Approximation as Strategic Communication

Manfred Krifka
Humboldt Universität Berlin
Zentrum für Allgemeine Sprachwissenschaft (ZAS) Berlin

West Coast Conference of Formal Linguistics
Berkeley, April 27 - 29, 2007

Failed attempts to introduce the metric system
Road signs in American National Parks
(Unfortunately, no photograph available, cited from memory)

Eagle Pass
7 miles
(11.265 kilometers)

From the land of bankers and watchmakers

Street sign in Kloten, Switzerland.

Overview of the talk:
1. Precise/Approximate Interpretations of number words:
   The basic phenomenon
2. Selection of optimal expressions and interpretations
   with preference for short expressions, approximate interpretations,
   and bidirectional OT: Krifka 2002
3. Selection of optimal expressions and interpretations
   with preference for short expressions and strategic communication:
   Krifka 2006
4. Selection of optimal scales
5. Principles for the construction of optimal scales
6. Adaptation of scales to language use:
   shortening of common expressions
7. Adaptation of language use to scales:
   under-use of complex expressions
8. A new perspective on language and language use
Overview of the talk:

1. Precise/Approximate Interpretations of number words: The basic phenomenon
2. Selection of optimal expressions and interpretations with preference for short expressions, approximate interpretations, and bidirectional OT: Krifka 2002
3. Selection of optimal expressions and interpretations with preference for short expressions and strategic communication: Krifka 2006
4. Selection of optimal scales
5. Principles for the construction of optimal scales
6. Adaptation of scales to language use: shortening of common expressions
7. Adaptation of language use to scales: under-use of complex expressions
8. A new theoretical perspective on language use

Precise vs. Approximate Interpretations of Number Words: The basic phenomenon

There were forty people at the workshop.

Precise interpretation: n = 40
Approximate interpretation: n = 40, e.g. 35 < n < 45

The context may favor particular precision levels (informal talk vs. report to funding agency)
but in many cases choice of precision level is left open within a given range.

Precision level can be indicated explicitly with particles (hedges):

*There were exactly forty people at the conference.*
*There were about forty people at the conference.*

...but also implicitly by choice of number words:

*There were forty people at the workshop.* => approximate interpretation
*There were thirty-eight people at the workshop.* => precise interpretation

*The line is forty centimeters long.*
*The line is forty point zero centimeters long.*

How does speaker infer precision level by choice of number words?
(This should depend on general pragmatic rules of language use.)

Explanation by general pragmatic principles

(Krifka 2002: ‘Be brief and vague! And how bidirectional OT allows for verbosity and precision’)

1. A preference for simple expressions
   cf. G. K. Zipf 1929, Relative Frequency: Principle of least effort
   A. Martinet 1962: Speaker economy
   L. Horn 1984: R-Principle
   St. Levinson 2000, Presumptive Meanings: I-Principle
   *forty > thirty-eight* (where a > b: ‘a preferred over b’)

2. A preference for approximate interpretations
   cf. P. Duhem 1904, balance between precision an certainty;
   E. Ochs Keenan 1976, vagueness helps to save face;
   reduction of cognitive effort with approximate interpretations, evident e.g. design of measuring devices in airplanes
   *approximate > precise*

3. Interaction of the two principles following Bidirectional Optimality Theory
   Blutner 2000 ‘Some aspects of optimality in natural language interpretation’
   Jäger 2002, ‘Some notes on the formal properties of bidirectional OT’
Optimal expression-interpretation pairs

3. Interaction of the two principles following Weak Bidirectional OT:
An expression-interpretation pair \((F, M)\) is optimal iff
there are no other optimal pairs \((F', M)\) or \((F, M')\)
such that \((F', M) > (F, M)\) or \((F, M') > (F, M)\).

Bidirectional Optimization and M-Implicatures

A Bi-OT account of M-Implicatures (Levinson 2000):
- Unmarked Expressions -- Unmarked Meanings;
- Marked Expressions -- Marked Meanings.

Many, many examples:
- Black Bart killed the sheriff.
- Black Bart caused the sheriff to die. (McCawley 1978)
- Mary went to the library and copied an article.
- Mary went to the library, and in addition copied an article.
- John went to school. / John went to the school.
- match box / box with matches

General explanation by two preferences, one for form, one for interpretation:
- Short expression > Long expression
- Stereotypical interpretation > Non-stereotypical interpretation

In the case at hand we have the following preference for interpretation:
- Approximate interpretation > Precise interpretation

Is it justified to assume a general preference for approximate interpretation,
as assumed in Krifka (2002)?

Overview of the talk:
1. Precise/Approximate Interpretations of number words:
   The basic phenomenon
2. Selection of optimal expressions and interpretations
   with preference for short expressions, approximate interpretations,
   and bidirectional OT: Krifka 2002
3. Selection of optimal expressions and interpretations
   with preference for short expressions and strategic communication:
   Krifka 2006
4. Selection of optimal scales
5. Principles for the construction of optimal scales
6. Adaptation of scales to language use:
   shortening of common expressions
7. Adaptation of language use to scales:
   under-use of complex expressions
8. A new theoretical perspective on language use

Explanation by Principles of Strategic Communication
cf. Krifka 2006, ‘Approximate interpretation of number words: A case for strategic communication’

Basic assumptions:
- Preference for short expressions, as before.
- No general preference for approximate interpretations
- Strategic communication,
  Parikh 2000, Communication, meaning and interpretation;
  Benz, Jäger, van Rooij (eds.) 2006, Games and Pragmatics, Oxford: Palgrave

General principle:
-- For addressee: Assume the speaker intends the most likely interpretation,
given the choice of expressions and the a-priori likelihood of message
-- For speaker: Assume the addressee selects the most likely interpretation,
given the choice of expressions and the a-priori likelihood of message

Basic question for interpretation:
Given the a-priori-likelihood of the communicated information
and general interpretation strategies,
what is the most likely interpretation of an ambiguous or vague form?
**Background: Modelling Precise and Approximate Interpretations**

Approximate interpretations of various granularity levels; ways of implementation:

1. Precise: numbers, / approximate: intervals:
   \[ [[\text{thirty-eight}]]_{\text{prec}} = 38, \quad [[\text{thirty-eight}]]_{\text{appr}} = [34..42] \]
2. Precise and approximate: intervals:
   \[ [[\text{thirty-eight}]]_{\text{prec}} = [38..38], \quad [[\text{thirty-eight}]]_{\text{appr}} = [34..42] \]
3. Various levels of precision: \([n]_{\text{lim}} = [n \pm \text{n/m}], \) e.g. \([\text{thirty-eight}]]_{1/10} = [38 \pm 3.8] = [34.2 .. 41.8] \)
   Precise interpretation as borderline case:
   \[ [[\text{thirty-eight}]]_{1/0} = [38 \pm 0] = [38..38] \]
4. Various levels of precision, using standard deviation as acceptable levels of fit:

\[ [[\text{thirty-eight}]]_{1/10} = (38 \pm 3.8) \]
\[ [[\text{thirty-eight}]]_{1/20} = (38 \pm 1.6) \]

**Strategic Communication:**

**Explanation of Precise/Approximate Interpretations**

Assumptions:

- General tendency for short expressions provided that they are interpretatively equivalent.
- No general preference for precise or approximate interpretations:
  - e.g. two interpretations, prec and appr, with \( p(\text{prec}) = p(\text{appr}) = 0.5 \)
- Within a given range, the a-priori probability of values is comparable or equal, e.g. \( p(38) = p(39) = p(40) = p(41) = p(42) = 0.02 \)
- Values that differ minimally (under discriminatory threshold, within acceptable level of fit) are interpretatively equivalent:
  - e.g. \( [[\text{thirty-eight}]]_{1/10} = [[\text{forty}]]_{1/10} \)
  - as \( [[\text{thirty-eight}]]_{1/10} = (38 \pm 3.8) = [34.2 .. 41.8] \)
  - \( [[\text{forty}]]_{1/10} = (40 \pm 4) = [36 .. 44] \)
  - and \( 40 \in [34.2 .. 41.8] \)
    and \( 38 \in [36 .. 44] \)

Precise values are within each other’s ranges.

**Strategic Interpretation:**

**Explanation of Precise/Approximate Interpretations**

\[ p(\text{prec}) = 0.5 \]
\[ p(\text{appr}) = 0.5 \]

Total probability:
- \( p(38) = 0.02 \)
- \( p(40) = 0.02 \)
- \( p(36 .. 44) = 0.16 \)

Winning interpretation:
- Total probability
- Least complex: \( \text{forty} \)
- \( 40 \pm 4 \)
- \( 38 \pm 3.8 \)
- \( 37 \pm 3.2 \)
- \( 38 \pm 3.8 \)
- \( 38 \pm 1.6 \)

Reported with least complex number within range?

- \( 38 \pm 3.8 \)
- \( 38 \pm 1.6 \)
- \( 38 \pm 0 \)
- \( 38 .. 38 \)

Winning interpretation:
- \( 38 \pm 3.8 \)
- \( 38 .. 38 \)
- \( 38 \pm 0 \)
- \( 38 .. 38 \)

 uttered
- \( \text{forty} \)
- \( 40 \pm 4 \)
- \( 38 \pm 3.8 \)
Overview of the talk:

1. Precise/Approximate Interpretations of number words:
   The basic phenomenon
2. Selection of optimal expressions and interpretations
   with preference for short expressions, approximate interpretations,
   and bidirectional OT: Krifka 2002
3. Selection of optimal expressions and interpretations
   with preference for short expressions and strategic communication:
   Krifka 2005
4. Selection of optimal scales
5. Principles for the construction of optimal scales
6. Adaptation of scales to language use:
   shortening of common expressions
7. Adaptation of language use to scales:
   under-use of complex expressions
8. A new theoretical perspective on language use

Length of expression is not always decisive

Preference for short expressions cannot explain all interpretation preferences:

- I did the job in twenty-four hours.    approximate
- I did the job in twenty-three hours.  precise
- I did the job in twenty-five hours.   approximate
- The house was built in twelve months. approximate
- The house was built in eleven months. precise
- The house was built in thirteen months. approximate
- Two dozen bandits attacked him.      precise
- Twenty-four bandits attacked him.    approximate

... and sometimes even makes the wrong predictions:

- Mary waited for forty-five minutes. approximate
- Mary waited for forty minutes.      precise
- I turned one hundred and eighty degrees. approximate
- I turned two hundred degrees.       precise
- Her child is eighteen months.       approximate
- Her child is twenty months.         precise
- John owns one hundred sheep.       approximate
- John owns ninety sheep.            precise

Alternative theory: A preference for simple, coarse-grained representations?

Coarse vs. fine grained levels of representation


Example: Minute scales of different granularity

<table>
<thead>
<tr>
<th>Value</th>
<th>Scale a</th>
<th>Scale b</th>
<th>Scale c</th>
<th>Scale d</th>
<th>Scale e</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>✔️</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>23</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>24</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>25</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>26</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>27</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>28</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>29</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td>30</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
</tbody>
</table>

Options for implementation of Coarse/Fine-Grained Scales

1. Values of measure functions are intervals, not numbers, but are named by numbers; for example: measurement of a duration $d$ in minutes, granularity levels $a$, $b$, $c$, $d$, $e$

   $\text{min}_a(d) = 18$, named 18
   $\text{min}_b(d) = 17.5 \ldots 22.5$, named 20
   $\text{min}_c(d) = 15 \ldots 25$, named 20
   $\text{min}_d(d) = 7.5 \ldots 22.5$, named 15
   $\text{min}_e(d) = 15 \ldots 45$, named 30

2. Values of measure functions are numbers, but their range of values is constrained; durations are mapped to the number with the best fit.

   $\text{RANGE}_{\text{min}_a} = \{1, 2, 3, 4, \ldots\}$
   $\text{RANGE}_{\text{min}_b} = \{5, 10, 15, \ldots\}$
   $\text{RANGE}_{\text{min}_c} = \{10, 20, 30, \ldots\}$
   $\text{RANGE}_{\text{min}_d} = \{15, 30, 45, \ldots\}$
   $\text{RANGE}_{\text{min}_e} = \{30, 60, 90, \ldots\}$

A well-known fact: “Errors” may add up, i.e. additivity does not hold in a strict way:

   $\text{min}(x + y) = \text{min}(x) + \text{min}(y)$ is not always satisfied; e.g. if $x$ and $y$ are durations of 14 minutes each, then $x + y$ is a duration of 28 minutes, $\text{min}(x) = 10$, $\text{min}(y) = 10$, $\text{min}(x + y) = 30$, but $\text{min}(x) + \text{min}(y) = 20$ (!)
Reinterpretation of approximate interpretation: Selection of the coarsest level in which number word occurs

The train will arrive in fifteen minutes.
- a. 15
- b. 10
- c. 10
- d. 10
- e. 10

The train will arrive in twenty minutes.
- a. 20
- b. 15
- c. 15
- d. 15
- e. 15

The train will arrive in eighteen minutes.
- a. 18
- b. 15
- c. 15
- d. 15
- e. 15

Explanation of choice of coarseness level by strategic communication

Assume each coarseness level is selected with equal probability, p(1, 2, ...) = p(5, 10, ...) = p(15, 20, ...) = p(30, 60, ...) = 0.2

Assume equal a-priori probabilities of durations, p(15) = p(16) = ... = p(20) = 0.01

Coarse scales and simple expressions: An evolutionary perspective on brevity?

Coarse-grained scales have expressions of reduced complexity (cf. Mrklić 2002)

Example: Complexity of number words on scales, average number of syllables
- a. one, two, three, four, ... one hundred: 273/100 = 2.73
- b. one, five, ten, fifteen, ... one hundred: 46/20 = 2.3
- c. one, ten, twenty, thirty, ... one hundred: 21/10 = 2.1

Suspection:
Scales develop in a way to enable complexity-based optimization, expressions of coarse-grained scales tend to be simpler.

Overview of the talk:
1. Precise/Approximate Interpretations of number words: The basic phenomenon
2. Selection of optimal expressions and interpretations with preference for short expressions, approximate interpretations, and bidirectional OT: Mrklić 2002
3. Selection of optimal expressions and interpretations with preference for short expressions and strategic communication: Mrklić 2005
4. Selection of optimal scales
5. Principles for the construction of optimal scales
6. Adaptation of scales to language use: shortening of common expressions
7. Adaptation of language use to scales: use of under-complex expressions
8. A new theoretical perspective on language use
An evolutionary perspective on brevity?
The optimization of scales.

Scales and hierarchies of scales of different granularity have to satisfy certain requirements to be useful for communication:

1. Requirement for scales: Equidistant of units (additive, sometimes logarithmic, cf. decibel; deci-/milli-/nano-/pico-, kilo-/mega-/giga-/tera-)
2. Requirements for scale hierarchies of different granularity: Scales of increasing granularity $S_i$, $S_{i+1}$ should increase granularity by the same factor, example: powers of 10:

- [10, 20, 30, 40, 50, 60, ...]
- [100, 200, 300, 400, 500, ...]
- [1000, 2000, 3000, 4000, ...]

The most natural step:

- decrease granularity by factor 1/2:

- $[1/2, 1, 1/2, 2, ...]$
- $[1/4, 1/2, 1/4, 1, ...]$

- cf. hour scale:

- $[1h, 2h, ...]$
- $[30min, 1h, 1h30min, ...]$
- $[15min, 30min, 45min, 60min, ...]$

- cf. J. Hobbs (2001) ‘Half Orders of Magnitude’, for a refined view of natural magnitude hierarchies with shifted half points, e.g. $[1...3...10...30...100...]$

Overview of the talk:

1. Precise/Approximate Interpretations of number words:

   - The basic phenomenon

2. Selection of optimal expressions and interpretations with preference for short expressions, approximate interpretations, and bidirectional OT: Krifka 2002

3. Selection of optimal expressions and interpretations with preference for short expressions and strategic communication: Krifka 2005

4. Selection of optimal scales

5. Principles for the construction of optimal scales

6. Adaptation of scales to language use: shortening of common expressions

7. Adaption of language use to scales: under-use of complex expressions

8. A new theoretical perspective on language use

An evolutionary perspective on brevity?
The optimization of scales

The expressions of values of scales align with the optimization of scales

Example: Expression of half points between powers of ten

Roman number writing (also motivated iconically, by shape of hand)

<table>
<thead>
<tr>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
<th>V</th>
<th>VI</th>
<th>VII</th>
<th>VIII</th>
<th>IX</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>X</td>
<td>XX</td>
<td>XXX</td>
<td>XL</td>
<td>L</td>
<td>LX</td>
<td>LXX</td>
<td>LXXX</td>
<td>XC</td>
<td>C</td>
</tr>
<tr>
<td>X</td>
<td>XX</td>
<td>XXX</td>
<td>XL</td>
<td>L</td>
<td>LX</td>
<td>LXX</td>
<td>LXXX</td>
<td>XC</td>
<td>C</td>
</tr>
<tr>
<td>C</td>
<td>CC</td>
<td>CCC</td>
<td>CD</td>
<td>D</td>
<td>DX</td>
<td>DXX</td>
<td>DXXX</td>
<td>XM</td>
<td>M</td>
</tr>
</tbody>
</table>

Simplification of word number ‘five’:

- English: fifteen (*fifteen), fifty (*fifty):

  - loss of diphthong, shortening
  - OEng: fif as word vs. filt- as prefix; vowel shift only affected i: (> ai), this did not affect nine in nineteenth

- Phonological simplification of ‘fifteen’, ‘fifty’ in colloquial German:

  - fuftzehn, fuftzig vs. regular fünfzehn, fünfzig:

  - unrounding ü > u, loss of n, shortening (3 to 2 morae)

- Phonological simplification of ‘half’ in German anderthalb ‘one and a half’, lit. ‘the second half’ vs. regular eineinhalb

- Morphological simplification of ‘fifty’ in Danish:

  - halvtrede vs. older halvredsinde-s-tve, similarly for ‘seventy’, ‘ninety’

The optimization of scales in language change

Development of number system in Old World history:

Babylonian system: based on 60; 60 and multiples thereof as estimation number e.g. in Ancient Greek (K. Menninger, 1962, Number words and number symbols):

The swineherd always sent them the best one of the fattened pigs for them to feast on, and the number of the swine remaining was only three hundred and sixty. (Odyssey)

Different complexity of number words in ancient languages; break after “60” has disappeared in modern languages

<table>
<thead>
<tr>
<th>Greek</th>
<th>Anglo-Saxon</th>
<th>Gothic</th>
<th>Celtic</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>pentē-konta</td>
<td>fiftig</td>
<td>coi-ca</td>
</tr>
<tr>
<td>60</td>
<td>hexē-konta</td>
<td>sixtig</td>
<td>ses-ca</td>
</tr>
<tr>
<td>70</td>
<td>hebdomē-konta</td>
<td>hund-seofontig</td>
<td>sibunt-e-hunt</td>
</tr>
<tr>
<td>80</td>
<td>ogdō-konta</td>
<td>hund-ahtatig</td>
<td>ahtaūt-e-hund</td>
</tr>
</tbody>
</table>

Phonological realizations of number words in current-day Russian:

Evidence for preferred reference points: Frequency of number words

If fine-grained / coarse-grained scales are used to report measurements, and if with coarse-grained scales, only certain number words occur, then these number words should occur more likely in a natural linguistic corpus containing measurement reports.


- Between 10 and 100, the powers of ten occur most frequently
- Frequency decreases with higher powers of 10, but there is a local maximum for 50
- Between 10 and 20, there are local maxima at 15, also at 12 ("dozen")

Example: Number words in British National Corpus, after H. Hammarström (2004), Properties of lower numerals and their explanation... (ms.)

Evidence for preferred reference points: Frequency of number words

Still an effect of complexity of expression?

In vigesimal number systems, '50' is more complex than '40'/60'

Question:

Is '50' nevertheless used as approximate number word?

Conflict cognitive preference / communicative preference

Cf. Hammarström (2004), Number bases, frequencies and lengths cross-linguistically.

Inspired by this work: Investigation of occurrences of number words on Norwegian vs. Danish web sites of Google (March 4, 2005)

<table>
<thead>
<tr>
<th>Number</th>
<th>Norwegian</th>
<th>Occurrences</th>
<th>Danish</th>
<th>Occurrences</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>tiue</td>
<td>61300</td>
<td>tyre</td>
<td>121000</td>
</tr>
<tr>
<td>30</td>
<td>treli</td>
<td>43700</td>
<td>tres</td>
<td>25400</td>
</tr>
<tr>
<td>40</td>
<td>farti</td>
<td>39200</td>
<td>fyre</td>
<td>26800</td>
</tr>
<tr>
<td>50</td>
<td>femti</td>
<td>81200</td>
<td>halvtreds</td>
<td>15500</td>
</tr>
<tr>
<td>60</td>
<td>seksli</td>
<td>19400</td>
<td>tres</td>
<td>36400</td>
</tr>
<tr>
<td>70</td>
<td>syfti</td>
<td>10200</td>
<td>halvjerds</td>
<td>581</td>
</tr>
<tr>
<td>80</td>
<td>atli</td>
<td>13100</td>
<td>fir s</td>
<td>3740</td>
</tr>
<tr>
<td>90</td>
<td>retti</td>
<td>13500</td>
<td>halvlems</td>
<td>540</td>
</tr>
</tbody>
</table>
Experiment on number estimation:
Decimal vs. vigesimal systems

Problem of corpus analysis:
There might be a wide variety of reasons for number word frequency
Hence: Experimental evidence is necessary.

Number estimation experiment:
Subjects should estimate number of dots on a screen;
all numbers between 12 and 100 were shown once
in nearly random order;
(two subsequent pictures differed by at least 15 points).

Hypothesis:
Speakers of languages with a vigesimal system
use complex multiples of 10 less often,
(i.e. number words for 50, 70, 90)
than speakers of languages with decimal system.

Complexity of multiples of 10 in Danish, Basque, Georgian

<table>
<thead>
<tr>
<th></th>
<th>Danish</th>
<th>Basque</th>
<th>Georgian</th>
</tr>
</thead>
<tbody>
<tr>
<td>tis</td>
<td>hamar</td>
<td>a1</td>
<td></td>
</tr>
<tr>
<td>fyve</td>
<td>hoge1</td>
<td>otsi</td>
<td></td>
</tr>
<tr>
<td>tredive</td>
<td>hoge1 ta hamar</td>
<td>otsdaati</td>
<td></td>
</tr>
<tr>
<td>fyvre</td>
<td>berroge1</td>
<td>ormotsi</td>
<td></td>
</tr>
<tr>
<td>halvtreds</td>
<td>berroge1 ta hamar</td>
<td>ormotsdaati</td>
<td></td>
</tr>
<tr>
<td>tres</td>
<td>hiruroge1</td>
<td>samotsi</td>
<td></td>
</tr>
<tr>
<td>halvijerds</td>
<td>hiruroge1 ta hamar</td>
<td>samotsdaati</td>
<td></td>
</tr>
<tr>
<td>firs</td>
<td>larroge1</td>
<td>otkhmo1</td>
<td></td>
</tr>
<tr>
<td>halvems</td>
<td>larroge1 ta hamar</td>
<td>otkhmo1</td>
<td></td>
</tr>
<tr>
<td>hundredre</td>
<td>ehun</td>
<td>asi</td>
<td></td>
</tr>
</tbody>
</table>

Examples of test items

![Graph showing number words in Norwegian and Danish Google Websites]
Description of experiment:

Number of dots in these examples: 79 - 30 - 73
Principles of presentation of items:
• Dots are spaced out regularly to allow for numerosity-based estimations (cf. Stanislas Dehaene 1997, The number sense),
• Dots were presented for 5 seconds,
• Every number between 12 and 100 occurred once in pseudo-randomized order,
• Subjects reacted by pronouncing a number.
• Instructions suggested to subjects to give estimations like forty, eighty,
  and changed midway through the experiment to force multiples of 10.
Experiment still ongoing:
  Norwegian (20) -- thanks to Anton Benz
  Danish (20) -- thanks to Torgrim Solstad
  Georgian (12) -- thanks to Rusudan Asatiani
  (Basque (7) -- thanks to Miren Lou Oinenderra;
    a more comprehensive experiment on Basque to be carried out soon.)
Number estimation experiment: Discussion

No valley at “50” for Danish/Georgian speakers:
- This scale position has high cognitive prominence, regardless of complexity (at least for the population the subjects are drawn from -- university students)

Explanation of peak at 70 with Norwegian (and German) speakers:
1. “50” is a natural estimation number, cf. St. Dehaene 1997, The number sense
2. Differences of +/- 5 are difficult to recognize for point sets around size 50; differences of +/- 10 are recognizable.
3. As “50” is a natural attractor, “60” is a less natural attractor, as it is not sufficiently distinct from “50”
4. “70” is a better attractor, as it is sufficiently distinct from “50”

Explanation of lack of peak for “70” for Danish and Georgian speakers:
- Even though “70” is a natural attractor for cognitive reasons, the complexity of the number word for “70” mitigates against it.

Explanation of peak at “40” with Danish/Georgian: Simplicity of number word?
- More experiments are needed,
- Basque seems particularly attractive due to greatest complexity differences.
Overview of the talk:

1. Precise/Approximate Interpretations of number words:
   - The basic phenomenon
2. Selection of optimal expressions and interpretations with preference for short expressions, approximate interpretations, and bidirectional OT: Krifka 2002
3. Selection of optimal expressions and interpretations with preference for short expressions and strategic communication: Krifka 2005
4. Selection of optimal scales
5. Principles for the construction of optimal scales
6. Adaptation of scales to language use: shortening of common expressions
7. Adaption of language use to scales: under-use of complex expressions
8. A new theoretical perspective on language use

Problems of the experiment

- Influence of second native language, e.g. Basque: Spanish/French, Welsh: English
- Influence of formal education in school, better don’t take university students as subjects!
- Make subjects estimate higher numbers: Run experiment with numbers n, 30 < n < 90

A new perspective on language and language use

Common functionalist belief:
“Grammars do best what speakers do most” (DuBois 1987)

But perhaps we also have to assume:
- Sometimes speakers do most what grammars do best,
- or perhaps:
- Speakers do least what grammars do worst.

That is:
- Meanings that are difficult to encode are expressed more rarely.

Related to Codability Hypothesis, Brown & Lenneberg (1954):
- Color words and color recognition.

Other evidence for this?
- For example, marking of aspectual distinctions,
- Progressives in English vs. German,
- difference in complexity, and corresponding frequency of use.

For example:

- John read the book.
- John was_aux reading the book.
- John las das Buch.
- John las gerade_au, das Buch.
- John war das Buch am_n, Lesen_nom (Rhenish dialects)

Overview of the talk:

1. Precise/Approximate Interpretations of number words:
   - The basic phenomenon
2. Selection of optimal expressions and interpretations with preference for short expressions, approximate interpretations, and bidirectional OT: Krifka 2002
3. Selection of optimal expressions and interpretations with preference for short expressions and strategic communication: Krifka 2005
4. Selection of optimal scales
5. Principles for the construction of optimal scales
6. Adaptation of scales to language use: shortening of common expressions
7. Adaption of language use to scales: under-use of complex expressions
8. A new theoretical perspective on language use

A new perspective on language use

An analogy to economics: Optimization of transaction costs
Example (unfortunately, not based on any study):
- Typical prices of goods at Dutch fleamarkets before January 1, 2002:
  - 2.5 guilders (ca. 1.136 Euros) or multiples thereof (Rijksdaalder)
- Typical prices of goods at Dutch fleamarkets after January 1, 2002:
  - 1 Euro (2.2 guilders) or multiples thereof

1 Euro (2.2 guilders) or multiples thereof
The End

1. Precise/Approximate Interpretations of number words:
The basic phenomenon
2. Selection of optimal expressions and interpretations
   with preference for short expressions, approximate interpretations,
   and bidirectional OT: Krifka 2002
3. Selection of optimal expressions and interpretations
   with preference for short expressions and strategic communication:
   Krifka 2005
4. Selection of optimal scales
5. Principles for the construction of optimal scales
6. Adaptation of scales to language use:
   shortening of common expressions
7. Adaptation of language use to scales:
   under-use of complex expressions
8. A new theoretical perspective on language use

Slides can be downloaded at: http://amor.rz.hu-berlin.de/~h2816i3x, Talks

Bibliography

Journal of Semantics 17, 189-216.
Dehaene, Stanislas, Mehler, Jacques (1992), “Cross-linguistic regularities in the frequency of
number words”, Cognition 43.
Dehaene, Stanislas (1997), The number sense. How the mind creates mathematics, Oxford
University Press, Oxford.
Duhem, Paul (1904), La theory physique, son objet et sa structure, Paris.
Hammarström, Harald (2004), “Number bases, length and frequency cross-linguistically”.
Horng, Laurence R. (1984), “Toward a new taxonomy for pragmatic inference: Q-based and R-
based implicature”, in D. Schiffrin, Meaning, form, and use in context: Linguistic
SKY - Yearbook of the Linguistic Association of Finland.
Jäger, Gerhard (2002), “Some notes on the formal properties of bidirectional Optimality
numerical expressions”, Journal of quantitative linguistics 8, 187-201.

Krifka, Manfred (2002), “Be brief and vague! And how Bidirectional Optimality Theory allows
for verbosity and precision”, in David Ristle & Dietmar Zaefifer, Sounds and systems:
Studies in structure and change. A Festschrift for Theo Vennemann, Berlin / New York,
Mouton de Gruyter, 439-458

Krifka, Manfred (2006), “Approximate interpretations of number words: A case for strategic
communication.”, in Gerlou Bouna, Irene Krämer & Joost Zwarts, Cognitive foundations of
interpretation, Amsterdam, Royal Netherlands Academy of Arts and Sciences, 111-126.
McCawley, James D. (1978), “Conversational implicature and the lexicon”, in Peter Cole,
Menninger, Karl (1962), Number words and number symbols, Dover.
Ochs, Elinor (1976), “The universality of conversational implicature”, Language in
Society 5, 67-80.
14, 473-514.
Parikh, Prashant (2000), “Communication, meaning, and interpretation.”, Linguistics and
philosophy. 23, 185-212.
Zipf, George K. (1929), “Relative frequency as a determinant of phonetic change”, Harvard
Studies in Classical Philology 40.