DATA TO HELP US FORM IMAGES: THE ABSTRACT AND NON-IMAGEABLE

Anita Peti-Stantić
University of Zagreb
OVERVIEW

1. INTRODUCTION AND THEORETICAL BACKGROUND

2. CROATIAN PSYCHOLINGUISTIC DATABASE
   - http://megahr.ffzg.unizg.hr/

3. METHODOLOGY AND GOALS

4. CONCLUSION
‘chair/a chair’
‘dog’
‘peace’
‘experience’
THREE MAIN QUESTIONS

1. WHAT ARE CONCRETENESS AND IMAGEABILITY?

2. WHY DO THEY MATTER?

3. HOW THEY CAN BE ACCESSED?
WHAT?

+ Concretness and imageability are **cognitive and psycholinguistic constructs**

WHY?

+ because they do not exist in an outside world beyond our head
+ they are our individual conceptualizations of the world
+ (similar to most abstract words)
CONCRETENESS AND IMAGEABILITY IN MENTAL GRAMMAR

+ MENTAL GRAMMAR - a complete collection of patterns, templates or schemas of the language stored in the brain/mind of a language user

+ CONCRETENESS
  + relation btw. lexical representation and the world (objects in the world)
  + EXPERIENTIAL

+ IMAGEABILITY
  + representation of a concept in our mental lexicon
  + participants asked to externalize their internal lexical representations
WHY?

+ COGNITIVE ADVANTAGE IN PROCESSING
  + AND MEMORY
+ CONCRETENESS EFFECT
  + BUT ALSO SALIENCE
COGNITIVE ADVANTAGE

FOR A LONG TIME, TWO INFLUENTIAL ACCOUNTS


MORE RECENTLY – DIFFERENT APPROACHES

HOW?

+ INDIRECT ACCESS THROUGH PSYCHOLINGUISTIC EXPERIMENTS

+ manipulating/combining
  + syntagmatic and syntactic

+ identification (lexical decision tasks, word-naming tasks, RT)

+ understanding (density of word-nets, semantic relations between words)

+ memorizing (short- and long-term memory, longer and better)
PSYCHOLOGY vs. LINGUISTICS
PSYCHOLOGY vs. LINGUISTICS

- Words are never used *in absentia*
  - Yet we test them one by one

- Sentence grammaticality and textual/discoursive coherence depends on the combinatorial potential of semantic and grammatical frames.
PSYCHOLOGY
processing

- LEXICAL PROCESSING
  - Concrete words (words referring to concrete concepts)
    - identified faster (lexical decision tasks, word-naming tasks, RT)
    - retained longer (and better, for combinatorial purposes) in long and short term memory

- BUT...
  - data from aphasia and ageing brain – reported both effects
  - Vigliocco et al. reported faster processing of abstract words
QUESTIONS to be asked

- Do syntagmatic relations of concrete and abstract words differ?
- If they do differ, how?
  - Sketch engine
- What does that mean for the language capacity?
THE BUILDING BLOCKS OF CROATIAN MENTAL GRAMMAR: CONSTRAINTS OF INFORMATION STRUCTURE
COLLABORATORS AND EXPERTS
CROATIAN PSYCHOLINGUISTIC DATABASE
https://doi.org/10.17234/megahr.2019.hpb

- Corpus HrWac – words more frequent than 3000 occurrences: 7695 Nouns, 3124 Adjectives, 2849 Verbs (random choice of 1000 of each)

- Psycholinguistic measures
  - concreteness, imageability
  - relative frequency, AoA

- 3000 words (1000 Nouns, Verbs and Adjectives)

- Additional 3000 words
  - 1500 most frequent words from the Frequency Dictionary of Croatian language
  - 1500 words from textbooks (Croatian Language, History, Mathematics, Natural Science, Geography)

- 2 sets of questionnaires – adults and elementary-school students
MEGACRO
WHY THE DATABASE? (1)

+ we wanted to produce a database for Croatian comparable to other databases:

MEGACRO
WHY THE DATABASE? (2)

1. RESEARCH INTERESTS

- comparability of data points with other databases
- lexicalization and conceptualization
- the relation of lexical network and syntax
- i.e. specificity of abstract words in languages such as Croatian, with suffixes –nje (gledanje [watching], trčanje [running]), -ost (urednost [tidiness], umjetnost [art]), -ota (dobrota [kindness], ljepota [beauty]), -stvo (zdravstvo [health (system)], ustrojstvo [structure])
MEGACRO
WHY THE DATABASE? (3)

+ 2. PRACTICAL INTEREST
  + relation between academic vocabulary and reading comprehension
MAIN LINGUISTIC RESEARCH QUESTIONS

+ Which features are universal and which are language-specific? Why?
  + mostly of typological, but also of ecological interest (how do the features of one language translate to another one)

+ How do these features affect the combinatorial potential of Nouns, Verbs and Adjectives?
  + at the interface of semantics and syntax
  + at the interface of semantics and phonology
CROATIAN PSYCHOLINGUISTIC DATABASE

+ METHODOLOGICAL GOALS

+ Which part of the language knowledge is common to all speakers of the language and which part is subject to variability?
  + adults vs. children (elementary school students)
  + two groups of adults (familiarity effect)

+ NEXT STEP - testing adults significantly different in their level of education
RELATION BETWEEN CONCRETENESS AND IMAGEABILITY: RESULTS

- All results calculated with Pearson correlation coefficient $r = 0.82$ ($p < 0.01$)
- Divided by the central value (3.17) into two groups - statistically different correlations for abstract ($r = 0.454$ ($p < 0.01$)) and concrete words ($r = 0.735$ ($p < 0.01$))
- As expected - less concrete, less correlated
SEPARATED PARTS OF SPEECH

Further analysis shows that the correlation is the lowest for **adjectives** ($r = 0.74; p < .01$), somewhat higher for **verbs** ($r = 0.82; p < .01$) and the highest for **nouns** ($r = 0.86; p < .01$)

COGNITIVE INCENTIVE - QUALITATIVE ANALYSIS NEEDED!
RELATION BETWEEN CONCRETENESS AND IMAGEABILITY: OPEN ISSUES

- high correlation in general, but...

- Examples in our database: mišljenje [thinking] (Concr = 1,27, Imag = 4,17), reinkarnacija [reincarnation] (Concr = 1,33, Imag = 4,07), obećanje [promise] (Concr = 1,7, Imag = 4,47), tajanstven [mysterious] (Concr = 1,67, Imag = 4,3), zadesiti [befall] (Concr = 1,93, Imag = 4,63)

- AVENUES OPEN FOR INTERPRETATION:
  - Two or more different meanings? Embodiment? Internal body-related sensory experiences? Emotional valence?
100 MOST CONCRETE WORDS IN THE CROATIAN DATABASE

+ SAME: all highly concrete
  + MOSTLY NOUNS (96%) !!!

+ EXCEPT:
  + differ substantially in corpus and relative frequency
    + telefon ‘phone’ – f 86.99 p/m !!
    + termostat ‘thermostat’ – f 2.53 p/m
    + plašt ‘robe’ – rf 2.41
    + olovka ‘pencil’ – rf 4.76

+ SIMILAR:
  + in AoA – approx. 4 - 10
100 MOST ABSTRACT WORDS IN THE CROATIAN DATABASE

- **SAME:** all highly abstract
- **EXCEPT:** Adjectives, Verbs and Nouns

**AND:**
- differ substantially in corpus and relative frequency
  - *nesuđen* ‘not meant to be’ f 2.47 p/m
  - *mišljenje* f 185.51 p/m
  - *posvemašnji* ‘widespread’ rf 1.48
  - *mišljenje* ‘thinking’ rf 4.76

- **SIMILAR:**
  - in AoA – approx. 7.5 – 14
CONCRETE vs. ABSTRACT

Concrete

Abstract

Noun  Verb  Adjective

32  19  49
CONCLUSION: CONCRETE AND ABSTRACT

- 100 most concrete lexemes in our database are mostly Nouns
- 100 least concrete lexemes in our database are Adjectives, Nouns and Verbs
- No conclusive evidence about the types of relations, but the density and composition of syntactic and semantic networks for concrete and abstract lexemes differ.