WORD-FORMATION AND
THE ORGANIZATION OF
THE MENTAL LEXICON

MEi:CogSci, WS 2016/17
UNIVERSITY OF VIENNA
29 NOVEMBER 2016
What is word-formation?
What is word-formation?

- Word-formation is part of morphology
What is word-formation?

- Word-formation is part of morphology
- Morphology consists of the following components:
What is word-formation?

- Word-formation is part of morphology
- Morphology consists of the following components:
  - **Word-formation** (production of new words)
    - Derivation, e.g. lehren ‘to teach’ → Lehrer ‘teacher’
What is word-formation?

- Word-formation is part of morphology

- Morphology consists of the following components:
  - **Word-formation** (production of new words)
  - Derivation, e.g. lehren ‘to teach’ → Lehrer ‘teacher’
    (combination of morphemes)
What is word-formation?

- Word-formation is part of morphology

- Morphology consists of the following components:
  - **Word-formation** (production of new words)
    - Derivation, e.g. lehren ‘to teach’ → Lehrer ‘teacher’
      (combination of morphemes)
    - Compounding, e.g. Lehrerzimmer ‘teachers’ room’
What is word-formation?

- Word-formation is part of morphology

- Morphology consists of the following components:
  - **Word-formation** (production of new words)
    - Derivation, e.g. lehren ‘to teach’ → Lehrer ‘teacher’ (combination of morphemes)
    - Compounding, e.g. Lehrerzimmer ‘teachers’ room’ (combination of words)
What is word-formation?

- Word-formation is part of morphology

- Morphology consists of the following components:
  - **Word-formation** (production of new words)
    - Derivation, e.g. lehren ‘to teach’ → Lehrer ‘teacher’
      (combination of morphemes)
    - Compounding, e.g. Lehrerzimmer ‘teachers’ room’
      (combination of words)
  - **Inflection** (production of word-forms)
What is word-formation?

- Word-formation is part of morphology
- Morphology consists of the following components:
  - **Word-formation** (production of new words)
    - Derivation, e.g. lehren ‘to teach’ → Lehrer ‘teacher’
      (combination of morphemes)
    - Compounding, e.g. Lehrerzimmer ‘teachers’ room’
      (combination of words)
  - **Inflection** (production of word-forms)
    E.g. lehre, lehrst, lehrt, etc.
What is word-formation?

- **Word-formation** is part of morphology.

- **Morphology** consists of the following components:
  - **Word-formation** (production of new words)
    - Derivation, e.g. lehren ‘to teach’ → Lehrer ‘teacher’
      (combination of morphemes)
    - Compounding, e.g. Lehrerzimmer ‘teachers’ room’
      (combination of words)
  - **Inflection** (production of word-forms)
    E.g. lehre, lehrst, lehrt, etc.
    Lehrbuch, Lehrbücher, etc.
Word-formation techniques

There are **five basic morphological techniques**:
Word-formation techniques

There are five basic morphological techniques:

- **Addition**, e.g. to teach → teach-er
There are five basic morphological techniques:

- **Addition**, e.g. to teach $\rightarrow$ teach-er
  - -er is an affix, more precisely a suffix
Word-formation techniques

There are five basic morphological techniques:

- **Addition**, e.g. to teach $\rightarrow$ teach-er
  - -er is an affix, more precisely a suffix
- **Substitution**, e.g. Marx-ism $\rightarrow$ Marx-ist
There are five basic morphological techniques:

- **Addition**, e.g. to teach $\rightarrow$ teach-er
  - -er is an affix, more precisely a suffix

- **Substitution**, e.g. Marx-ism $\rightarrow$ Marx-ist

- **Modification**, to impórt $\rightarrow$ impórt
There are five basic morphological techniques:

- **Addition**, e.g. to teach → teach-er
  - -er is an affix, more precisely a suffix
- **Substitution**, e.g. Marx-ism → Marx-ist
- **Modification**, to impórt → ímport
- **Conversion**, to cut → a cut
Word-formation techniques

There are five basic morphological techniques:

- **Addition**, e.g. to teach → teach-er
  -er is an affix, more precisely a suffix
- **Substitution**, e.g. Marx-ism → Marx-ist
- **Modification**, to impórt → impórt
- **Conversion**, to cut → a cut
- **Subtraction**, e.g. Russian biologija ‘biology’ → biolog ‘biologist’
Word-formation techniques

There are five basic morphological techniques:

- **Addition**, e.g. to teach → teach-er
  - -er is an affix, more precisely a suffix
- **Substitution**, e.g. Marx-ism → Marx-ist
- **Modification**, to impórt → ímpor
t- **Conversion**, to cut → a cut
- **Subtraction**, e.g. Russian biologija ‘biology’ → biolog ‘biologist’

These techniques represent all possible cognitive operations that can be performed on a morphological form.

Suffixation and suffix ordering

real →
Suffixation and suffix ordering

real → real + -ize
Suffixation and suffix ordering

real $\rightarrow$ real + -ize $\rightarrow$
$\rightarrow$ real + -iz + -ation
Suffixation and suffix ordering

real $\rightarrow$ real + -ize $\rightarrow$

$\rightarrow$ real + -iz + -ation $\rightarrow$

$\rightarrow$ real + -iz + -ation + -al
Suffixation and suffix ordering

real $\rightarrow$ real + -ize $\rightarrow$
$\rightarrow$ real + -iz + -ation $\rightarrow$
$\rightarrow$ real + -iz + -ation + -al

- Note that an alternative ordering of the suffixes is not possible, i.e. *real-iz-al-ation, *real-al-ation-ize, etc. do not exist.
Explanation of the order of the suffixes

- According to the type of information used in suffix ordering:
  1) phonological
  2) morphological
  3) syntactic
  4) semantic
  5) statistical
  6) psycholinguistic
  7) cognitive
  8) templatic

Manova & Aronoff (2010)
The mental lexicon

- A notion used in linguistics and psycholinguistics
  - Psycholinguistics studies how language works in the brain
- The mental lexicon is something like a mental dictionary where systematic information about language (words and their use) is stored in an easily accessible way
- There are different opinions about what information exactly is stored in the mental lexicon
  - Some linguists believe that only whole words (and no suffixes) are represented in the mental lexicon
Structure of the talk

- My research: Languages analyzed
- My cognitive approach
- Hypotheses about the organization of the mental lexicon
- Two psycholinguistic experiments
- Discussion of the results of the experiments
- Conclusions about what is stored in the mental lexicon
My research: Languages analyzed

- Slavic
  - Bulgarian (South Slavic)
  - Russian (East Slavic)
  - Polish (West Slavic)
- Germanic
  - English
  - German
- Romance
  - Italian

- Editor of papers on about 30 typologically diverse languages
Slavic word versus English word

**Slavic word**

(PREFIX)-BASE-(DERIVATIONAL SUFF)-(THEMATIC MARKER)-(INFLECTIONAL SUFF)

non-evaluative  evaluative

**English word**

(PREFIX)-BASE-(DERIVATIONAL SUFF)-(THEMATIC MARKER)-(INFLECTIONAL SUFF)

non-evaluative

Ø
The combinability of the English suffix *-ist*

<table>
<thead>
<tr>
<th>SUFF1</th>
<th>Lexical category of SUFF1</th>
<th>Followed by SUFF2</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ist</td>
<td>N</td>
<td>-dom, -ic, -y, -ize</td>
</tr>
</tbody>
</table>

Data from Aronoff & Fuhrhop (2002), based on OED, CD 1994
Explanation of the combinability of -\textit{ist} 

According to the type of information used in suffix ordering:

1) phonological
2) morphological
3) syntactic
4) semantic
5) statistical
6) psycholinguistic
7) cognitive
8) templatic

Manova & Aronoff (2010)
My approach

According to the type of information used in suffix ordering:

1) phonological
2) morphological
3) syntactic
4) semantic
5) statistical
6) psycholinguistic
7) cognitive
8) templatic

Manova & Aronoff (2010)
English -ist: My cognitive approach

<table>
<thead>
<tr>
<th>SUFF1</th>
<th>Lexical category of SUFF1</th>
<th>SUFF2</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ist</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>N: -dom (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ADJ: -ic (631), -y (5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>V: -ize (3)</td>
</tr>
</tbody>
</table>

Table from Manova (2011)
Data from Aronoff & Fuhrhop (2002), based on OED, CD 1994

Nouns, adjectives and verbs are seen as cognitive categories, cf. Langacker (1987).
Lexical categories:
Noun (N), Adjective (ADJ) and Verb (V)

Langacker (1987), based on *relationality* (i.e. +/- relational) and *way of scanning* (whether summarily scanned, i.e. conceived statistically and holistically, or sequentially scanned, i.e. mentally scanned through time), recognizes *things* (N), *processes* (V) and *modifiers* (ADJ).
-ist: Fixed combinations

<table>
<thead>
<tr>
<th>SUFF1</th>
<th>Syntactic category of SUFF1</th>
<th>SUFF2</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ist</td>
<td>N</td>
<td>N: -dom (2)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ADJ: -ic (631), -y (5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>V: -ize (3)</td>
</tr>
</tbody>
</table>

Table from Manova (2011)
Data from Aronoff & Fuhrhop (2002)
Types of SUFF1-SUFF2 combination

- **Fixed (unique)**
  - SUFF1 combines with only one particular SUFF2 of a major lexical category, N, V, ADJ
-ist: Predictable combinations

<table>
<thead>
<tr>
<th>SUFF1</th>
<th>Syntactic category of SUFF1</th>
<th>SUFF2</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ist</td>
<td>N</td>
<td>N: -dom (2) ADJ: -ic (631), -y (5) V: -ize (3)</td>
</tr>
</tbody>
</table>

Table from Manova (2011)  
Data from Aronoff & Fuhrhop (2002)
Types of SUFF1-SUFF2 combination

- **Fixed (unique)**
  - SUFF1 combines with only one particular SUFF2 of a major lexical category, N, V, ADJ

- **Predictable**
  - SUFF2 applies by default – the majority of words are derived by that suffix.
    - Suffixes that ‘compete’ with the default suffix are unproductive and derive no more than 10 words
Hypotheses

- **H1**: If SUFF1 tends to combine with only one SUFF2 of a major lexical category (N, ADJ, V), SUFF1-SUFF2 combinations are unique pieces of structure and speakers should know them by heart.

- **H2**: If speakers know suffix combinations by heart, they should be able to differentiate between existing and non-existing combinations and existing combinations should be recognised with higher accuracy and faster than non-existing ones.
Experiment 1

- **Participants:** 64 native speakers of Polish
  - age: M=23.2, SD=1.76
  - no history of developmental dyslexia or reading disabilities
  - non-linguists

- **Materials:** 60 items
  - 30 existing suffix combinations from Polish, e.g.:
    - -ar-nia as in kawi-ar-nia ‘café’
  - 30 non-existing suffix combinations created by changing the order of the suffixes of the legal ones or by manipulating phonemes, e.g.:
    - from the existing -ar-nia → -ni-ar or -ur-nia.
  - 2 lists
    - each with the suffixes of the other in reverse order
    - each participant saw all combinations
Experiment 1: Procedure

- **Task:** decide as quickly and as accurately as possible if a combination exists or not
- **Training:** a few examples of derivations of existing and non-existing words with two suffixes in Polish to ensure that the participant understands the task
- **List of items:** participants received a list of existing and non-existing suffix combinations and have to complete the task
- **Maximum time for decision:** 10 minutes
Experiment 1: Accuracy of recognition of existing and non-existing combinations

- Acc for existing: \( M=81.72\% \), \( SD=0.29 \)
- Acc. for non-existing: \( M=75.99\% \), \( SD=0.22 \)

The result is statistically significant: 
\[ t(63)=2.34, \quad p=0.02 \]
Experiment 2

- **Participants**: 53 native speakers of Polish
  - age: $M=21.43$, $SD=1.83$
  - no history of developmental dyslexia or reading disabilities
  - non-linguists

- **Task**: Press the right arrow button if a string of letters is an existing combination or the left CTRL button if it is not. In case of a doubt, behave as if a stimulus does not exist.

- **Materials**: 88 items, randomized with the E-prime 2.0 software
  - 44 existing and 44 non-existing suffix combinations
  - non-existing combinations produced as in Experiment 1
  - 2 lists
    - each with the suffixes of the other in reverse order
    - each participant saw all combinations
Experiment 2: Procedure
Experiment 2: Accuracy

Existing combinations:
\[ M_{ACC} = 81\%, \ SD = .09 \]

Non-existing combinations:
\[ M_{ACC} = 74\%, \ SD = .12 \]

The result is statistically significant:
\[ t(52) = 3.03, \ p = 0.004 \]
Experiment 2: RTs

Existing combinations:
1333 ms
$M_{RT}=1333.14$, $SD=420.57$

Non-existing combinations:
1610 ms
$M_{RT}=1610.38$, $SD=556.02$

The difference is statistically significant:
$t(51)=-7.53$, $p<0.001$
Experiment 2: Mean accuracy of the productive combinations (derive > 10 words)

Productive combinations:
$M_{ACC} = 86\%, \ SD=.09$

Unproductive combinations:
$M_{ACC} = 75\%, \ SD=.11$

The difference is statistically significant:
$t(51)=7.81, \ p<0.001$
Experiment 2: Mean RTs of the productive combinations (derive > 10 words)

Productive combinations: $M_{RT}=1288.44$, $SD=429.14$

Unproductive combinations: $M_{RT}=1421.01$, $SD=488.41$

The difference is statistically significant: $t(51)=-4.08$, $p<0.001$
The results of the two experiments converge:

- The accuracy of recognition of the existing combinations is significantly higher than the accuracy of recognition of the non-existing combinations.
- The reaction times to the existing combinations are significantly shorter than to the non-existing ones.
- Thus, recognition of suffix combinations seems to resemble recognition of words and non-words in psycholinguistics.
- The productive combinations are recognized more accurately and faster than the unproductive combinations.
Suffixation and suffix ordering

real $\rightarrow$ real + -ize $\rightarrow$

$\rightarrow$ real + -iz + -ation $\rightarrow$

$\rightarrow$ real + -iz + -ation + -al
Suffixation in the mental lexicon: Conclusions

real → real + -ize →
→ real + -iz + -ation →
→ real + -iz + -ation + -al

- Our research shows that parts of words such as suffixes and suffix combinations are stored in the mental lexicon.
- When speakers produce complex words, they, most probably, do not attach suffixes step by step but use them as wholes, i.e. as -ization, -ational and, maybe, -izational.
Thank you!

stela.manova@univie.ac.at
http://homepage.univie.ac.at/stela.manova/
Selected references


