What is in a morpheme? Theoretical, experimental and computational approaches to the relation of meaning and form in morphology

Workshop introduction

Stela Manova, Harald Hammarström, Itamar Kastner
University of Vienna / Uppsala University / Humboldt University of Berlin

i. Two paradoxes

- **Paradigm Function Morphology (PFM), Construction Morphology (CxM), A-morphous morphology (A-morphM)**

<table>
<thead>
<tr>
<th>There are no morphemes BUT there is morphology.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Inflectional) morphemes are listed as markings (exponents) without meaning in the lexicon. Thus, e.g., PFM defines form and meaning based on the paradigm function (PF): PF(L,σ) = (R,σ) (Stewart &amp; Stump 2007). The PF value of a paradigm cell (L,σ) of the lexeme (L) is the pairing of this cell’s realization R with the morphosyntactic property set σ.</td>
</tr>
</tbody>
</table>

- **Distributed Morphology (DM)**

<table>
<thead>
<tr>
<th>There is no morphology by itself BUT there are abstract morphemes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>DM relies on syntactic structures and ‘morpheme’ is an abstract unit that refers to a syntactic terminal node and its content, not to the phonological expression of that terminal. Morphology is distributed between syntax and phonology. Vocabulary items are exponents that relate form and meaning but they insert late (post syntactically).</td>
</tr>
</tbody>
</table>

ii. Our view

<table>
<thead>
<tr>
<th>There is morphology and there are morphemes.</th>
</tr>
</thead>
<tbody>
<tr>
<td>There are three possible ways to approach the relation of meaning and form in morphology:</td>
</tr>
<tr>
<td>A. Form and meaning emerge simultaneously</td>
</tr>
<tr>
<td>B. The association is from meaning to form</td>
</tr>
<tr>
<td>C. The association is from form to meaning</td>
</tr>
</tbody>
</table>

The most important difference between these scenarios consists in the fact that in scenarios B and C meaning may be assigned at the level of word, i.e. one may claim that morphemes do not have meaning of their own or even that there are no morphemes at all (in scenario B).

Language is a positional system.

- With respect to meaning, (forms of) morphemes can be evaluated:
  1) in isolation (as building blocks of morphology, e.g. English -s, -en, -ed, -er)
  2) based on their position in the word form (i.e. templatically, e.g. inflection is outside derivation; prefixes, suffixes, infixes, interfixes are also established positionally)
  3) based on combinations with other morphemes, e.g. E. writ-er-s but not *small-er-s, thus two different -er suffixes, one that derives nouns (agents) and one that expresses comparative degree of adjectives
Morphemes associate form and meaning, like in scenario A, but this association is not trivial and involves scenarios B and C at the different stages of derivation and in comprehension and production.

iii. The structure of the talk

1. Morphemes and exponents: What is the difference?
2. Why are there no morphemes that relate meaning and form?
3. It is hard to write about morphology without morphemes
4. Basic elements and rules: The importance of having a clear definition of a morpheme
5. Meaning and form in computer science (the language C++)
6. Morphemes in psycholinguistics
7. Morphemes in computational linguistics
8. The workshop papers

1. Morphemes and exponents: What is the difference?

1.1. Definitions from Understanding Morphology (Haspelmath & Sims 2010)
   - **morpheme**: the smallest meaningful part of a linguistic expression that can be identified by segmentation; a frequently occurring subtype of morphological pattern
   - **exponent**: when a morphological pattern (e.g. -ed) expresses an inflectional feature value (e.g. past tense), it is the exponent of that feature value

1.2. Definitions from What is Morphology? (Aronoff & Fudeman 2011)
   - **morpheme**: a word or a meaningful piece of a word that cannot be divided into smaller meaningful parts. Examples include school, read, or the re- and -ing of rereading.
   - **exponent**: the marker of a given morphosyntactic feature. For example, [s] is the exponent of plural in the word kits

2. Why are there no morphemes that relate meaning and form?

2.1. The answer

<table>
<thead>
<tr>
<th>PFM, CxM, A-morphM &amp; DM: The relationship between meaning and form is not perfect, i.e. not one-to-one.</th>
</tr>
</thead>
</table>

(1) **Bulgarian -ta:**
   - meče 'little bear'
   - meče-ta 'little bear-plural, i.e. little bears'
   - meče-ta-ta 'little bear-plural-definite, i.e. the little bear'

(2) **English -er:**
   - writ-er (AGENT)
   - (bottle) open-er (INSTRUMENT)
   - strong-er (COMPARATIVE)

   - What should -ta in Bulgarian and -er in English mean?
(3) Noun plural
a. **Bulgarian**
   kniga 'book' – PL knig-i
   more 'sea' – PL more-ta
   etc.

b. **English**
   book – PL book-s
   ox – PL ox-en
   etc.

❖ Which morpheme signals plural in Bulgarian and English?

2.2. **The solution**

➢ EITHER meaning is assigned only at the level of word (construction in CxM) and there are no morphemes (cf. theories such as PFM, CxM, A-morphM)
➢ OR there are morphemes but they are abstract (cf. DM)

3. **It is hard to write about morphology without morphemes**

3.1. **PFM**

(4) Second-person imperative active forms of the 9th-conjugation verb KRĪ 'buy' in Sanskrit (forms extracted from Table 5, Bonami & Stump 2017)

   2sg    krī-ṇī-hi
   2du    krī-ṇī-tam
   2pl    krī-ṇī-ta

“In Sanskrit, for example, the default expression of membership in the 9th conjugation is a suffix -ni (sandhi form -ṇi), as in the imperative form krī-ṇi-ta 'you (pl.) buy!', and the default expression of second person singular subject agreement in active imperatives is -hi, as in krī-ṇi-hi 'you (sg.) buy!'.” (Bonami & Stump 2017, our emphasis here and in (5))

(5) Rules of exponence in Sanskrit (see (20) in Bonami & Stump 2017)

   a. I, Xv[9th conjugation], {} → Xni
   b. II, Xv, {2sg imp active} → Xhi

❖ Pay attention to what is in {}.

3.2. **CxM**

“The suffix -ig expresses the notion ‘somewhat’, that is, it has a relativizing meaning.” (Booij 2007: 40, our emphasis)

3.3. **DM**

(6) Vocabulary of English (fragment), from Bobaljik (2015)

   a. [PAST] ↔ -t /v_; where V ∈{dream, dwell etc.}
   b. [PAST] ↔ ∅ /v_; where V ∈{run, hit, fly etc.}
   c. [PAST] ↔ -d /v_

❖ Note the bidirectional arrows.
4. Basic elements and rules: The importance of having a clear definition of a morpheme

4.1. Mathematics

- There are basic elements such as different types of numbers (natural, rational, irrational, real, etc.)
  - **Natural numbers**: all positive integers (whole numbers) and the zero.
  - **Rational numbers**: all numbers that can be expressed as the fraction of two integers.
  - **Irrational numbers**: all real numbers that are not rational.
  - **Real numbers**: all rational and irrational numbers, i.e. any point anywhere on the number line.

- There are operations (=rules) such as addition, subtraction, multiplication, etc. that manipulate those numbers
- Numbers and operations are defined from the very beginning

  ❖ Numbers (correspond to morphemes in our case) and operations (correspond to operations such as e.g. merge and agree in linguistics) cannot be redefined. Imagine what would happen if the definition of e.g. addition in mathematics changes or if there are arbitrary orderings of the operations in calculations such as (2+3)x4 + 5?

4.2. Chess (games in general)

- Basic elements in terms of a fixed number of figures
- Fixed rules
- Every player knows the figures and the rules from the very beginning

  ❖ The current situation in morphology can be described in the following way: one of the players assumes that rules depend on the chessboard (cf. PFM), the other player believes that the rules are encoded only in the figures (cf. DM).

5. Meaning and form in computer science (the language C++)

5.1. Roman numbers

(7) I, II, III, IV, V, etc.

- Each symbol always stands for the same value. All the “I” signs always represent the value 1 (one) wherever they are placed and the “V” sign always represents a value of 5 (five), etc.

Each symbol always represents the same value. The meaning of a symbol does not depend on its position, i.e. the system of the Roman numbers is not positional with respect to the values of its elements: “I” means 1(one) in IV (four) and in VI (six), although in IV it is subtracted while in VI it is added. Subtraction and addition are not elements but operations.

- Compare with semantic compositionality in linguistics

5.2. Decimal system

(8) Basic elements (digits): 0,1,2,3,4,5,6,7,8,9

1 This section is based on the following Internet resource: [http://wwwcplusplus.com/](http://wwwcplusplus.com/).
The same symbol does not always represent the same value. The value of a symbol depends on the position of that symbol, i.e. the decimal system is **positional**. The meaning of the symbol "1" in the decimal system: 1 (= “I” in Roman numbers); 10 (= X); 100 (= C), etc.

- 123 is not 1+2+3 but 100+20+3
- Compare with the understanding of iconicity and semantic compositionality in linguistics

### 5.3. Binary code

1. **Binary code**

(9) 110011111010010100

   - In C++, this binary expression can be represented either as (9), (10) or (11):

(10) 212628 (decimal, i.e. ten basic digits, as illustrated in (8))

(11) 0637224 (octal, i.e. eight different basic digits, from 0 to 7)

   - In C++ octal numbers always begin with a 0 digit
   - Cf. (5a): I, Xv[9th conjugation], {} \(\rightarrow\) Xnī

(12) 0x33e94 (hexadecimal, sixteen different basic digits, the numbers from 0 to 9 and the letters A, B, C, D, E and F)

   - In C++, hexadecimal numbers are preceded by 0x (zero, x).
   - Cf. (5a): I, Xv[9th conjugation], {} \(\rightarrow\) Xnī

Complex tasks and large amounts of information can be handled with a very limited number of basic elements BUT one MUST know those elements and the rules of the system they form from the very beginning. A complex system that operates with a limited number of basic elements relies on templates and assigns meaning not only to elements but also to positions.

### 5.4. Change of meaning in the decimal system and in linguistics

13. 12 \(\rightarrow\) 123  (the meaning of 12 changes due to addition of a slot)

   - Bg. žen-a ‘woman’ \(\rightarrow\) žen-a-ta ‘woman-DEF, i.e. the woman’

14. 12 \(\rightarrow\) 13  (the meaning of 12 changes due to substitution in an existing slot, cf. substitution of inflection)

   - Bg. žen-a \(\rightarrow\) žen-i ‘woman-PL, i.e. women’

15. 45 \(\rightarrow\) 456 \(\rightarrow\) 4566  (the meaning of 45 and 456 changes due to addition of slots)

   - Bg. more-e ‘sea’ \(\rightarrow\) more-ta ‘sea-PL, i.e. seas’ \(\rightarrow\) more-ta-ta ‘sea-PL-DEF, i.e. the seas’

**Is language a positional system? Yes, it is.**
A few **facts from morphology and syntax** in support of this conclusion (the list is incomplete):

- The differentiation between roots / stems and affixes is positional
- Stratal approach: Level 1 and Level 2 affixes are defined positionally
- Template morphology is entirely positionally defined
- Layered morphology (cf. semantic scope) is positionally defined
- Position classes in morphology
- There are positional restrictions on the placement of an affix in a word, cf. affix ordering constraints
- Movement in syntax
- Word order in syntax

Interesting to compare this to phonology, where similar positional systems can be found although no meaning as such is represented

Now recall 2. above “Why are there no morphemes that relate meaning and form?” and the PFM, CxM, A-morphM & DM answer to this question: The relationship between meaning and form is not perfect (i.e. not one-to-one)

| If language is a positional system, its form-meaning mappings cannot always be one-to-one because the meaning of an element in a positional system depends on the position of the element. |

6. **Morphemes in psycholinguistics**

6.1. **Lexical decision task (Meyer and Schwanefeldt 1971)**

- A cornerstone in the research on how polymorphemic words are processed
- Involves decisions at the level of word, that is, in terms of words versus non-words

6.2. **The "affix stripping" model (Taft and Forster 1975, 1976, Taft 1979)**

- Lexical decision task that involves **stem identification**: a prefixed word is accessed via its stem even when this stem is not a word in its own right, i.e. *unhook* is related to *hook* in the same way as *persuade* to *suade* and both the word *hook* and the non-word *suade* are stored in the mental lexicon; we need *suade* for the recognition of *persuade* and *dissuade*.
- Contemporary work corroborating to this model (Fruchter et al. 2013; Fruchter and Marantz 2015) demonstrates that speakers obligatorily decompose the (visual) stimulus into morphemes, look these up in the mental lexicon, and recombine them. All three steps can be individually observed and manipulated (show sensitivity to frequency, family density and so on).

6.3. **Affix identification**

- Affixes as parts of words are identified and obligatorily decomposed, even in masked priming and even if their forms are only orthographically identical to real affixes (Rastle et al. 2004; Stockall and Marantz 2006; Crepaldi et al. 2010, 2013; Lewis et al. 2011)
- Affixes can be identified and processed even in isolation or without having a contentful stem to attach to (Crepaldi et al. 2016; Lázaro et al. 2016; Beyersmann et al. 2016; Manova & Brzoza 2015)
6.4. Models with no morphemes
- Naive Discriminative Learning - NDL (Baayen et al. 2011, Plag & Balling 2016, see also the discussion of the model in Marantz 2013)
- NDL is explained in the next section

Experimental techniques used in the workshop studies
- Auditory primed lexical decision task (Creemers et al.)
- Phrase-completion task: phrases with missing inflection (Franzon et al.)
- Comprehension task: click on the right picture given a stimulus; and production task: picture naming (Kapatsinski & Harmon)

7. Morphemes in computational linguistics

7.1. Morpheme-based models
- Unsupervised Learning of Morphology (ULM)

7.2. A-morphous models
- Naïve Discriminative Learning (NDL)

7.3. Comparison of NDL and ULM
- See Appendix 1

8. The workshop papers

8.1. Three papers on experimental linguistics
Ava Creemers, Amy Goodwin Davies & Robert J. Wilder (University of Pennsylvania): Morphological priming of Dutch complex verbs is independent of semantic transparency
- Auditory primed lexical decision experiment which manipulates prime–target pairs with respect to their morphological, semantic, and phonological relatedness
- Investigation of the role of semantic transparency by teasing apart semantic and morphological effects
- Morphological structure is explicitly represented in memory and morphological identity is distinct from mere semantic and phonological similarity

Francesca Franzon, Rosa Rugani, Dunia Giomo & Chiara Zanini (University of Padova): Effects of animacy on the processing of morphological Number: a cognitive inheritance? A psycholinguistic study
- Language does not encode morphologically all possible information present in the referential world
- Focus on Number (encoded morphologically) and its intersection with animacy (when it is not encoded morphologically)
- A phrase-completion task: Noun phrases of a demonstrative and a noun appeared on the screen one at a time; the demonstrative or the noun lacked the inflectional morpheme
- How one looks for semantics determines what one finds

Vsevolod Kapatsinski & Zara Harmon (University of Oregon): Entrenchment in comprehension constrains semantic extension in production
- Comprehension task: click on the right picture given a suffixed noun
✓ Production task: name pictures using one of a set of suffixes
✓ Form-meaning mapping is not trivial but differs in comprehension and production: high frequency mapping causes semantic extension in production but entrenchedment in comprehension

8.2. Two papers on theoretical linguistics

Roland Pfau (University of Amsterdam): Morpheme Repair
✓ Production can be repaired morpheme-based, the repair is telling of the ordering of operations in morphology
✓ Derivational morphemes are inserted post-syntactically at the level of Morphological Structure
✓ Meaning-to-form mapping

Gereon Müller (University of Leipzig): Deponent Morphemes: A Case Study of Linkers in German Compounds
✓ Optimality-theoretic analysis
✓ The meaning (function) associated with a morpheme is not entirely fixed / can be discarded, which calls for a more flexible approach to the form-meaning association in the morpheme
✓ The interpretation of a morpheme changes with the change of the morpheme’s position in the word form

8.3. Fieldwork

Ekaterina Lyutikova & Sergei Tatevosov (Lomonosov Moscow State University): Doing form and meaning in a field: a few reflections on Buriat and Nenets
✓ Relates theory and practice
✓ A certain class of semantic generalizations can only be properly identified through their formal manifestation
✓ Fieldwork involves both meaning-to-form and form-to-meaning analyses of data

8.4. Morphomes
(purely morphological forms that cannot be defined in terms of meaning; since Aronoff 1994, morphomes are seen as the strongest evidence for the existence of morphology proper)

Borja Herce (University of Surrey & University of the Basque Country): Forms with(out) meaning: What can we learn from morphomes?
✓ ‘Morphomes’ or ‘meromorphomes’ are exponents with either no meaning of their own or with a meaning which can hardly be delimited
✓ Analysis from the perspective of Canonical Typology, i.e. the analysis is based on the definition of a canonical morpheme and deviations from the canon

8.5. Corpus-driven analysis with a typological orientation

Taras ZakhARKo (University of Zurich): Mining corpora for form-meaning associations: perspectives for corpus-driven typology
✓ What can we learn about the form-meaning relationship from large amounts of data?
✓ Promotes an information-theoretic approach
✓ How important is morphosyntactic information?
✓ The language signal encodes more information than what is commonly assumed
References


Marantz, Alec (2013). No escape from morphemes in morphological processing. *Language and Cognitive...*


## Appendix 1: A comparison of Naïve Discriminative Learning and Unsupervised Learning of Morphology

<table>
<thead>
<tr>
<th>Naïve Discriminative Learning (NDL)</th>
<th>Unsupervised Learning of Morphology (ULM)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Type of method</strong></td>
<td></td>
</tr>
<tr>
<td>A-morphous, psycholinguistic, computational</td>
<td>Morpheme-based, computational</td>
</tr>
<tr>
<td><strong>Data</strong></td>
<td></td>
</tr>
<tr>
<td>Prepared (semantically-annotated) corpora</td>
<td>Unannotated text corpora</td>
</tr>
<tr>
<td><strong>Units of analysis</strong></td>
<td></td>
</tr>
<tr>
<td>Words &amp; sounds in terms of bigrams (unigrams and trigrams could also be used)</td>
<td>Words in terms of strings of letters and morphemes (morphemes are substrings of letters)</td>
</tr>
<tr>
<td><strong>The role of morphemes</strong></td>
<td></td>
</tr>
<tr>
<td>None, HOWEVER, the annotation of the corpus is based on affixal semantics, s. ‘Outcomes’ in the next row which illustrates the learning of causative meaning</td>
<td>Repeating substrings of letters could be morphemes; after weighting of such substrings and their frequencies, the established morphemes are typically evaluated against a linguistic gold standard, e.g. in the MorphoChallenge, <a href="http://morpho.aalto.fi/events/morphochallenge/">http://morpho.aalto.fi/events/morphochallenge/</a>, i.e. substrings are compared with a list of existing morphemes prepared in collaboration with linguists</td>
</tr>
<tr>
<td><strong>Illustration</strong></td>
<td></td>
</tr>
<tr>
<td>(data from Table 1 in Plag &amp; Balling 2016)</td>
<td></td>
</tr>
<tr>
<td><strong>Words</strong></td>
<td><strong>Outcomes</strong></td>
</tr>
<tr>
<td>baptize</td>
<td>‘baptize’</td>
</tr>
<tr>
<td>modernize</td>
<td>‘modern’, ‘make’</td>
</tr>
<tr>
<td>optimize</td>
<td>‘optimal’, ‘make’</td>
</tr>
<tr>
<td>size</td>
<td>‘size’</td>
</tr>
<tr>
<td><strong>The role of semantics</strong></td>
<td></td>
</tr>
<tr>
<td>To learn the meaning of the causative -ize, one must prepare the corpus semantically, i.e. one must annotate the causative verbs modernize and optimize as ‘modern_make’ and ‘optimal_make’</td>
<td>None, i.e. semantics is not considered, though it could be, see the next row</td>
</tr>
<tr>
<td><strong>Questions:</strong></td>
<td></td>
</tr>
<tr>
<td>• If one annotated baptize and size as ‘bapt_make’ and ‘s_make’ respectively, would NDL learn that baptize and size were causative verbs?</td>
<td></td>
</tr>
<tr>
<td>• What would happen if one wrote in the annotation ‘modern_run’, ‘optimal_run’ or ‘modern_read’, ‘optimal_read’?</td>
<td></td>
</tr>
<tr>
<td><strong>Goals</strong></td>
<td></td>
</tr>
<tr>
<td>Learning morphology without morphemes; based on frequency distributions of words and weighting of associations between (bi)grams and outcomes, NDL might learn to generalize about -ize, i.e. to differentiate between causative and non-causative -ize in words that are not semantically annotated (see Baayen et al. 2011 for a detailed introduction to NDL)</td>
<td>Learning natural language’s morphology from unannotated text corpora; relies on comparison, grouping and weighting of substrings (of letters) and their frequencies (see Hammarström &amp; Borin 2011 for an overview of ULM research). Semantic representations of extracted form-based morphemes may also be inferred using the principle that semantically related morphemes tend to occur in similar contexts (e.g., as in Baroni et al. 2002).</td>
</tr>
</tbody>
</table>