2. SOME DEFINITIONS

Some Definitions

word: Turns out to be hard to define

- minimum free form (Lieber (2010:3), “[a form comprising] one or more morphemes that can stand alone in a language”): are the, and, my words? are they “free”? and what are “morphemes”

- items separated by spaces in writing: works for written languages and shows that speakers have a sense of words, but it doesn’t always work out cleanly for languages that do not have a standardized written tradition and even some that do! How many words are in the following list? wallpaper wall plug leapfrog leap year blacklist black belt

I was going to give samples of weird word breaking or none at all from transcriptions in languages without standard writing systems, but all the examples I could find did a pretty good job of making rational word breaks!

- non-separable linguistic units

  Q: [a clueless student] *What less is the student? A: *Clue.
  Aronoff & Fudeman (2005) cite a song from the musical Camelot:
  It’s May, it’s May, the month of “yes, you may,”
  The time for every frivolous whim, proper or im-,
  When all the world is brimming with fun, wholesome or un-

- linguistic units that have internal integrity

  That’s not at all possible.
  *That’s im-at all-possible.
  a very green lawn, a greener lawn
  *a very greenhouse, *a green-er-house
  Does the existence of fan-f…in’-tastic show that fantastic is not a word?

lexeme: - Lieber—“…[a family] of words that differ only in their grammatical endings or grammatical forms”
  Haspelmath—“a ‘dictionary word’…an abstract entity that can be thought of as a set of word forms”
  - Katamba—“abstract base for forms of the same word”
  - Aronoff & Fudeman—“…a word with a specific sound and a specific meaning [whose] shape may vary depending on syntactic context.”
  - Hayes—“an inflectable stem”

word-form: - Haspelmath—a “text word”, i.e. a concrete entity that “can be pronounced and used in texts—every word-form belongs to one lexeme”
  - Katamba—“a particular physical realisation of [a lexeme] in speech or writing”

sing, sings, singing, sang, sung are the word-forms that belong to the lexeme SING
citation form: a form of a lexeme chosen by convention as the name of the lexeme (for languages with published dictionaries, normally the head word for a dictionary entry)

SING is a lexeme whose citation form is “sing” and which is realized by the word-forms sing, sings, singing, sang, sung

paradigm: - Haspelmath and Aronoff & Fudeman—“the set of all the inflected forms that a lexeme assumes”
- Hayes—adds that the word-forms in a paradigm are “arranged according to the categories defined by the [system of inflectional features]”

<table>
<thead>
<tr>
<th>TENSE:</th>
<th>NUMBER:</th>
<th>NUMBER:</th>
</tr>
</thead>
<tbody>
<tr>
<td>person</td>
<td>sing.</td>
<td>plural</td>
</tr>
<tr>
<td>PERSON:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>ich</td>
<td>frage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>wir</td>
</tr>
<tr>
<td>2</td>
<td>du</td>
<td>fragst</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ihr</td>
</tr>
<tr>
<td>3</td>
<td>er/sie/es</td>
<td>fragt</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sie</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fragen</td>
</tr>
<tr>
<td>PERSON:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>ich</td>
<td>frage</td>
</tr>
<tr>
<td></td>
<td></td>
<td>wir</td>
</tr>
<tr>
<td>2</td>
<td>du</td>
<td>fragtest</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ihr</td>
</tr>
<tr>
<td>3</td>
<td>er/si/es</td>
<td>fragte</td>
</tr>
<tr>
<td></td>
<td></td>
<td>sie</td>
</tr>
<tr>
<td></td>
<td></td>
<td>fragten</td>
</tr>
</tbody>
</table>

word family (= lexeme family): [not an “official” piece of terminology, but useful] a set of lexemes that are related by having the same base (see below)

the word family of the lexeme SING includes also the lexemes RESING, SONG, SINGER, SINGABLE, UNSINGABLE, UNSUNG and, more marginally, SING-ALONG, SING-SONG, SONGFEST

morpheme: the smallest unit of meaning that cannot be further subdivided into meaningful units

undernourished comprises the morphemes under (“less than adequate”), nourish (“provide food”), ed (“in an affected state”); unnourished comprises the morphemes un (“not”) + the last two morphemes just mentioned. Although under includes the string of sounds un found in unnourished, our definition of “morpheme” does not allow us to divide under into two morphemes un+der because (i) the un of under does not provide the meaning “not” (or any other identifiable meaning), and (ii) der has no separate existence in any English lexeme. Parallel arguments can be applied in rejecting the ish of nourish as a separate morpheme even though it contains a substring of the phonemes found in foolish, which comprises the morphemes fool (“person without good sense”) and ish (“like a…”).

affix: a form attached to a base to form a new word-form (and potentially a new lexeme)

base: any form to which an affix is added (not everyone distinguishes base vs. stem, using “stem” in the sense of “base” as defined here)

stem: Haspelmath and Katamba—the base of an inflected word form

Lieber (2010:35): “The stem is usually the base that is left when inflectional endings are removed.”
The lexeme **SING** includes the (unaffixed) stem (or base) **sing**, to which the affix **-s** can be added to form the word-form **sings**. To the same base can be added the affix **-er** to form a new lexeme **singer**. The lexeme **SINGER**, in turn has the word-form **singer**, which can be the stem of the inflected form **singers** and the base of the derived form **non-singer**.

**Root:**
- Katamba—the irreducible core of a word, present in all the manifestations of a lexeme
- Haspelmath—“a base that cannot be analyzed any further into constituent morphemes” (roots and affixes are the primitive elements of morphology)

Deciding on what a root is will be language specific and time specific. There seems to be no problem in recognizing two roots in **cat’s-eye** referring to a type of precious stone, but **daisy** would be a root for all English speakers today, even though historically it is from **day’s eye**. **Hippopotamus** is a root in English even though it comprises two roots in Greek (**hippo-** and **pot-**). The fact that we can call this animal a **hippo** does not mean that /hippo/ is the “root” in English. Clearly, English **hippo** is created by shorting a long word even though, more or less by accident, the form corresponds to a root in the Greek original.
Practice With Terminology

A paradigm for the lexeme FRAGEN appears above. Here are some related items:

(ich) **frage** ‘(I) ask, (I) am asking’
(sie) **fragten** ‘(they) asked’
**Frage** ‘(a) question’
**Fragen** ‘questions’
**fraglich** ‘questionable’
**unfraglich** ‘unquestionable’
**fraglicher** ‘more questionable’
**fragliche** ‘questionable (fem.sing.)’, as in **fragliche Sache** ‘matter in question’
**fraglichere** ‘more questionable (fem.sing.)’

Using the terminology on the preceding pages…

(1) What would we call these words, taken as a group?

(2) What are the different **lexemes** represented by these words?

(3) Identify the following morphological elements in these words:
   - root
   - affix
   - base
   - stem

(4) What would be the likely **citation form** for ‘questionable’?

---

1 When used as direct modifiers of nouns, German adjectives are marked to agree with the nouns they modify in gender, number, and case. When used as predicate adjectives, however, adjectives are unmarked. Compare the sentence **diese Sache ist fraglich** ‘this matter is questionable’ with the example phrase.