

Chapter 12

Frequency effects in
morphology
From M. Haspelmath (2002),
Understanding Morphology



1

Contents of this chapter

- Word frequency: how often words are used
- Role in language (processing)
- Role in morphology:
 - Frequency in inflection
 - Direction of language change
 - Patterns of irregularity
 - Strength of synonymy blocking

2

Token/type frequency

- Token frequency: the number of times a given word form is used
 - e.g. the number of times the noun *bank* is used in a given *corpus* (=collection)
 - lemma frequency: e.g. the noun *bank* + *banks*
- Type frequency: the number of times a particular *type* of form is used
 - e.g. all past tenses in a given corpus
- HF = high-frequency LF = low-frequency
 - note: do not confuse with fundamental frequency, F0

3

Effects of word frequency

- Predictability: HF expressions are more predictable because their occurrence is more likely
- Memory strength: HF expressions are more easily remembered
- Fast retrieval: HF expressions are more quickly retrieved from memory (test on following slide!)

Some rules of grammar are influenced by word frequency

- relation grammar – frequency

4

bat

5

lush

6

Measuring frequency

- Frequency books, corpora
 - some freely available
 - Corpus Internet
- Ghits = "Google hits"
 - Bing hits (May 2017):
 - bat = 16.1 m
 - lush = 13.9 m

7

Inflectional categories

- Markedness :
 - passive = 'marked' compared to active
 - active sentence: no morpheme
 - passive sentence "special" morphology
 - plural = 'marked' compared to singular, etc.
- Markedness correlates exactly with frequency
 - active sentences more frequent than passives
- So... are some structures more frequent because they are unmarked, or are some structures unmarked because they are more frequent?
 - and then why are they more frequent?

8

Frequency and shortness

- HF words tend to be short in many languages (Zipf's Law)
 - one syllable; 2 or 3 segments
 - easily checked for English or French
 - short essay: check for Chinese
 - problems of definition: what is a word?
- Frequent inflectional categories may not be expressed at all
 - e.g. 1st ps sg present tense in verbs
 - nominative singular case in nouns
 - derivation of female from male professions

9

100 most frequent words in English

the be to of and a in that have I it for not on
with he as you do at this but his by from
they we say her she or an will my one all
would there their what so up out if about
who get which go me when make can like
time no just him know take person into
year your good some could them see
other than then now look only come its
over think also back after use two how our
work first well way even new want
because any these give day most us

10

"Frequency reversals"

- Some nouns normally appear in the plural (although they do have a singular), and for such nouns the plural is sometimes 'unmarked' compared to the singular
- Welsh
 - dail 'leaves' deilen 'leaf'
 - ffa 'beans' ffaen 'bean'
 - mefus 'strawberries' mefusen 'strawberry'
- Greater memory strength of frequent categories

11

Direction of change

- Historical change often advances into the direction of a frequent category
 - e.g. house [s] – houses [-z-]
 - some speakers change the plural into [-siz]
 - singular is the most frequent form
 - house 7.7b – houses 1.0b (Ghits)
 - voiceless also phonologically unmarked compared to voiced

12

Frequency and irregularity

- If there are irregularities in inflection, these affect the most frequent lexemes
 - already seen many examples, e.g. suppletion
 - Old English: four irregular verbs are 'be, do, go, want'
- fast retrieval leads to shorter forms
 - if you do something more often, you get more skilled and will do it more quickly (music, sports, cooking)
 - listener needs less information for recognition
- increased memory strength leads to less regularization

13

Recent research

- Li Yi'ou (SISU PhD 2015-)
- The productivity of tone sandhi rules in Qingdao Chinese
 - only three lexical tones; different tone sandhi rules (also producing T4) - which are productive?
 - (big) experiment with Q speakers
 - frequency factor: which environments for TS occur more often? ~productivity

14

Synonymy blocking and frequency

- Blocking: to sweep - a *sweep (broom)
 - The more frequent the blocking word is, the greater is its blocking strength
- | root | derivation | blocking word |
|----------|---------------|---------------------|
| big | ?bigness | size (very freq.) |
| warm | ?warmness | warmth |
| high | ?highness | height |
| true | ?trueness | truth (very freq.) |
| real | ?realness | reality |
| readable | ?readableness | readability (infr.) |
- (frequency: look up or use Google hits)

15

Summary

- Token & type frequency: relevant to morphology (and phonology) in many ways
- Closely linked to psycholinguistics: frequent forms are more quickly retrieved and more firmly stored
- Seen in inflectional categories and elsewhere
- Relevant to synonymy blocking
 - Various investigations possible

17

Thank you~~

Next week: optional class

Time for

- any questions
- essay help

18