ASSIGNMENT ABOUT ENGLISH, BASED ON EXPLORATORY EXERCISE CH. 4 (H&S)

Assignment II, Chapter 4, Lexicon

- This assignment must be handed in on Canvas, but it will NOT be graded – feedback will be provided in the form of group discussions (on methodology and results) and reports in class.
- If you want, you are free to prepare the assignments in groups, but everyone must hand in an individual report.
- Your report should be maximally 1000 words. It should contains your test items (targets and fillers) and their frequencies, some information about your participants, and your analysis/conclusion.

Intro/background:

In chapter 4, we argued that frequency has an effect on whether a complex word is directly given in the lexicon. But we did not ask two important morphological questions: if it is possible for a word-form like *insane* to be stored in the lexicon either via morphemes or directly, how can we know how the word is stored? And how can we know that frequency is an important factor? There are various methods for testing these questions. This exercise has two purposes: to introduce one of these methods, and to test the hypothesized relationship between frequency and lexical storage.

The basic methodology involves asking speakers of English to judge how related two morphologically related words (one basic, one derived) are to each other in meaning. The assumption underpinning this task is that speakers should judge complex words that are stored as morphemes as being closely related in meaning to their bases, because the basic and derived form share a lexical entry. By contrast, if a complex word is stored as its own lexical entry, we expect it to be judged as less close in meaning to its root. By manipulating the frequency of the complex word, we can test for a correlation between frequency and meaning closeness. Under the assumption that meaning closeness reflects the way of lexical storage, a correlation (or lack thereof) indicates whether frequency is a relevant factor.

In this exercise, we will use the English suffix *-ity* (as in *obscurity, immensity*, etc.).

Instructions:

- Step 1: Select 20 word pairs to be included in your experiment. A good way to do this, is to use a reverse dictionary (e.g. Muthmann 1999), which groups all words with a given suffix together. Use a frequency dictionary (e.g. Leech et al. 2001 for British English or Davies 2010 for American English) to gather information about the token frequency of these words.
  N.B. Please, consult these sources IN THE LIBRARY, so that your fellow students can also use them at all times!
  Alternatively, you can use the SUBTLEX-US database, which gives information about the frequency of American English words in a corpus of English subtitles. You can find the database through this link: http://crr.ugent.be/papers/SUBTLEX-US%20frequency%20list%20with%20PoS%20and%20Zipf%20information.zip. In the column FREQqount you find the total number of times a word is included in the corpus. The Zip-value indicates the frequencies of words (values 1-3 are low-frequency words, values 4-7 are high-frequency words).
  Sort your word pairs into 'frequent' and 'infrequent' groups. Remove words of intermediate frequency, in order to create two maximally distinct groups.
  Next, add a variety of 'filler' words with their roots; 5 per group. These fillers should be

derived lexemes that range in semantic similarity to their roots, but do not contain the suffix *-ity*. Finally, write each of the derived words (targets and fillers) and its root on a card, with the complex word first.

- Step 2: Based on the discussion in Section 4.3, develop specific predictions about the closeness judgements of your participants.
- Step 3: Run the experiment. Try to find at least five participants. Present them with the cards and ask them to rate how much the meaning of the first word is related to the meaning of the second word, on a scale from 1 to 7 (in which 7 means 'very closely related' and 1 means 'not at all related'). Write down the results per card and per speaker. Make sure you present target and filler items in mixed and random order.
  N.B. If you have trouble finding participants, you may try to recruit them on-line, via https://hinative.com/.
- Step 4: Analyze the data
- Step 5: Evaluate your predictions in light of your results. What do the data indicate about the relationship between frequency, semantic transparency and lexical storage?