**Class 7 (Week 3, R): Sideways interfaces II, phonology and processing**

**To do**

- Read **Zhang, Lai & Sailor 2011** for Thursday.

_____ will present Zhang & al.'s findings and interpretation

_____ will sketch out how the same hypotheses could have been tested using one or two other methods (that we'll discuss on Tuesday): artificial grammar learning, study of lexical/construction choice, priming...

- First **homework, about paradigms**, will be posted tonight. Due in 2 weeks (Oct. 29). This doesn't mean you should spend twice as much time on it as usual! Next HW will be computing exercise on Week 5 material, so no reason to make this one due any earlier.

**Overview**: What is grammar and what is processing?

## 1. Actually, that's way too big a question

- I've seen a lot of criteria proposed (or just used) that don't convince me:
  - if it's frequency-sensitive, it's not grammar
  - if it's variable, it's not grammar
  - if it's phonetically gradient, it's not grammar
- These all rely on a-priori assumptions about what grammar can be, but if we're trying to figure out what grammar is, we can't make those assumptions.
- Instead, in the first half of today let's look at some frequency effects and how they could fit in to our model of language; in the second half, we'll look at consequences of speech planning and lexical access occurring in real time.

## 2. Classic frequency effect: English irregular verbs

- There are only about 200 of them, but they are disproportionately likely to be frequent (e.g., Bybee & Slobin 1982).
- Top 25 most frequent verbs (Oxford English Corpus)—irregulars are in bold:

| | | | |
|---|---|---|---|
| **1. be** | **8. know** | **15. give** | **22. feel** |
| **2. have** | **9. take** | 16. use | 23. try |
| **3. do** | **10. see** | **17. find** | **24. leave** |
| **4. say** | **11. come** | **18. tell** | 25. call |
| **5. get** | **12. think** | 19. ask | |
| **6. make** | 13. look | 20. work | |
| **7. go** | 14. want | 21. seem | |

- Locus of explanation?

***Diachrony***

- In order to learn an irregular past tense form, you have to be exposed to it enough times → low-frequency verbs will tend to regularize from one generation to the next (*bode* > *bided*).
- Kirby 2001: simulation study

***Processing***

Dual-route model (see Pinker 2000 for overview and application to this case)

- When you want to say a past tense, there's a race between retrieving a stored form (which might be irregular) and creating the form via the *–ed* rule.
- The more frequent stored form → higher resting activation → more likely to win the race.
  - low-frequency verbs may get pronounced as regular, even if speaker knows irregular form.

***Grammar? (I don't think anyone has proposed it for this case, but it's a logical possibility)***

- Some constraints are sensitive to frequency.

| /bowd/, cf. [bajd] | I-O FAITH(hi freq) | O-O FAITH | I-O FAITH(lo freq) |
|---|---|---|---|
| bowd | | *! | |
| ☞ bajdɨd | | | * |

or split O-OFAITH by frequency.

- Or there's just one I-O FAITH constraint, but its *ranking* is a function of frequency

**3. Ng 2010: Singapore English prosodic boundaries**

- Singapore English has strong glottalization at prefix-stem, stem-stem, but not stem-suffix boundary
  - *mis-understand* [misʔɑndəstæn] , *stop-over* [stɔpʔovə], *magic-al* [mædʒikØəɯ](p. 8)
  - o Ng analyzes this in terms of p-word structure: let's sketch it out

- Tone pattern is roughly **L* (ˈM M*) H** (p. 11)
- Domain of tone assignment ≈ p-word
  - tone pattern generally re-starts in compounds: *century egg* (ˈMH)(ˌH) (p. 13)
  - tone pattern may or may not restart at prefix-stem boundary: *un-install* (ˌH)-(LˈH) ~ (L-(LˈH)) (p. 12)
  - tone doesn't restart at stem-suffix boundary: remove-able (LˈMMH) (p. 12)
- Initialisms show varying degrees of prosodic merger:

Society for the Prevention of Cruelty to Animals, Anglo-Chinese Junior College, National Registration Identity Card, National Trade Unions Congress (supermarket)

| | *Least merger* | | | *Most merger* |
|---|---|---|---|---|
| a. *SPCA* | ((((ˈH)ˌH)ˌH)ˌH) | | | |
| b. *ACJC* | ((((ˈH)ˌH)ˌH)ˌH) | (((ˈH)ˌH)ˌMH) | | |
| c. *NRIC* | ((((ˈH)ˌH)ˌH)ˌH) | (((ˈH)ˌH)ˌMH) | ((ˈH)ˌMMH) | |
| d. *NTUC* | ((((ˈH)ˌH)ˌH)ˌH) | (((ˈH)ˌH)ˌMH) | ((ˈH)ˌMMH) | (ˈMMMH) |

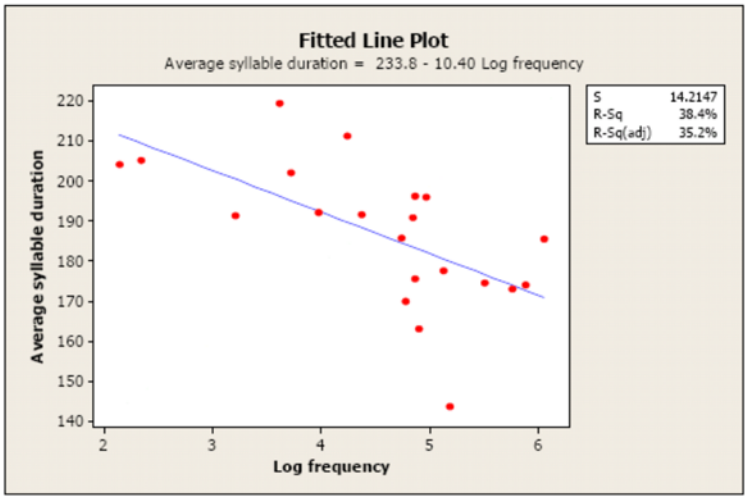
(p. 23)

- Ng finds correlation between which group an initialism belongs to and its # of Google hits.

***Why?***

- Frequency determines speed of production, perhaps because of faster access:

Figure 4: Frequency and duration




(p. 31)

- Constraints are then sensitive to speed, e.g. "Grammatical word accessed at speed *n* allows only *n* levels of stress"
  - Result is a prosodification of higher-frequency words that results in fewer stresses.
  - This is an interesting way of removing the need for the grammar to refer to frequency
    - Predicts that if we can manipulate speaking rate or retrieval speech independent of word frequency, we'll get similar effects.

(58) *Do not restore faithful stress to destressed initialisms*

| $[[M][O][E]]_{S_2}$ | WRAP | *SCLASH$_S$ | STRESS |
|---|---|---|---|
| ☞ a. (('M),OE) | | | * |
| b. (('M),O,E) | | W* | L |
| c. (('M)(,O)(,E)) | | W* | L |
| d. ('M)('O)('E) | W* | | L |

accessed at "speed 2" ($S_2$), so allows only two levels of stress (b and c have tertiary stresses)

(p. 33)

## 4. Hammond 1999: English rhythm rule

*thìrteen mén* or *thirtèen mén*?

- In survey, shift is more likely if adjective is more frequent: *nàive fríend* vs. *obèse chíld*
- Hammond proposes morpheme-specific faithfulness constraints, whose ranking depends on the word's frequency.

## 5. Löfstedt 2010: frequency-specific constraints

- Famous paradigm gaps in Swedish result when vowel shortening produces too much of a quality change.

| STEM | NEUTER | GLOSS |
|---|---|---|
| bl[oː$_T$] | bl[ɔ$_L$] + t: | 'blue' |
| v[iː$_T$]t | v[ɪ$_L$] + t: | 'white' |
| v[iː$_T$]d | v[ɪ$_L$] + t: | 'wide' |

quality change (from Tense to Lax) is not too big

(p. 152)

| STEM | NEUTER | GLOSS | ALLÉN (PL) | GOOGLE (-A) |
|---|---|---|---|---|
| gr[ɑː]d | INEFFABLE[51] | 'straight' | 0 | 7,140 |
| l[ɑː]t | INEFFABLE | 'lazy' | 0 | 581,000 |

quality change (would be from [ɑː] to [a]) is too big

(p. 154)

- But! Sufficiently frequent words don't have a gap

| | | | | |
|---|---|---|---|---|
| gl[ɑː]d | gl[a] + t: | 'happy' | 29 | 2,110,000 |

(p. 154)

- For each of the vowels that can show a gap, there seems to be a frequency cut-off above which there's no gap. (Löfstedt shows this for some phenomena in other languages too) E.g.,

| STEM | NEUTER | GLOSS | ALLÉN (PL) | GOOGLE (-A) |
|---|---|---|---|---|
| gr[ɑː]d | INEFFABLE[51] | 'straight' | 0 | 7,140 |
| l[ɑː]t | INEFFABLE | 'lazy' | 0 | 581,000 |
| gl[ɑː]d | gl[a] + t: | 'happy' | 29 | 2,110,000 |

frequency counts from different corpora

(p. 154)

- Löfstedt's solution: faithfulness constraints penalizing vowel changes are indexed to frequency:

| | /glad + t:/ 'happy' neut. Cf. [glɑːd] (*Freq* /glad/ = 2,110,000) | $\sigma_{\mu\mu}$ ↔ [+stress] | [+LONG] ↔ [+TENSE] | IDENT [Long C] /V_ | *MAP (ɑ,a) (7140) | *MAP (ɑ,a) (581,000) | M-PARSE | *MAP (ɑ,a) (2,110,000) |
|---|---|---|---|---|---|---|---|---|
| a. | glɑːtː | *! | | | | | | |
| b. | glɑt: | | *! | | | | | |
| c. | glɑːt | | | *! | | | | |
| d. > | glɑt: | | | | | | | * |
| e. | ⊙ | | | | | | *! | |

(p. 167)

## 6. Boersma 1999: lexical-access *constraints*

- The problem: in Dutch, you want to be able to recognize [rɑt] as either /rɑt/ or /rɑd/.
- If you try to use a standard grammar to map perceived form to underlying form, you'll always pick the faithful one:

(7) *Failure to recognize the wheel*

| [rɑt] | *VOICEDCODA | MAXVOI |
|---|---|---|
| *☞* \|rɑt\| 'rat' | | |
| \|rɑd\| 'wheel' | | * |

This is a comprehension tableau:
input = perceived phonetic form
output = lexical entry

(p. 4)

- So, Boersma proposes a family of constraints *LEX(*x*) "don't recognize any utterance as lexical item *x*" (one for each lexical item).

- Ranking depends on word's frequency:

(10) *A strong tendency to recognize the rat*

| [rɑt] | *LEX (\|rɑd\| 'wheel') | *VOICED CODA | MAXVOI | *LEX (\|rɑt\| 'rat') |
|---|---|---|---|---|
| ☞ \|rɑt\| 'rat' | | | | * |
| \|rɑd\| 'wheel' | *! | | * | |

(p. 5)

- Actually, it's a bit more complex: *LEX(*x*/context=*y*) to allow for semantic context to matter

## 7. More proposals in which grammar refers to frequency (if we have time)

Can we think of ways to determine whether grammar makes direct reference to frequency, or sees only to the outcome of lexical access?

- Coetzee 2008: a lexical item's frequency determines how likely it is to be assigned to a given lexical class on any production occasion
- Myers 2005: how can lenition be both postlexical and sensitive to lexical frequency?
  - proposes a diachronic solution, where high frequency results in a more lenited lexical entry over time (exemplars? see Pierrehumbert 2001), but plays no synchronic role
  - diachronic and synchronic explanations should make different predictions about effects of priming on production...
- Alcántara 1998 (English): high-frequency exceptions can be protected by high-ranking idiosyncratic constraints

- Carlson & Gerfen 2011 (not a proposal about grammar, but a cool case): when a Spanish diphthong loses stress (say, because of suffixation), it should monophthongize. But it' variable:

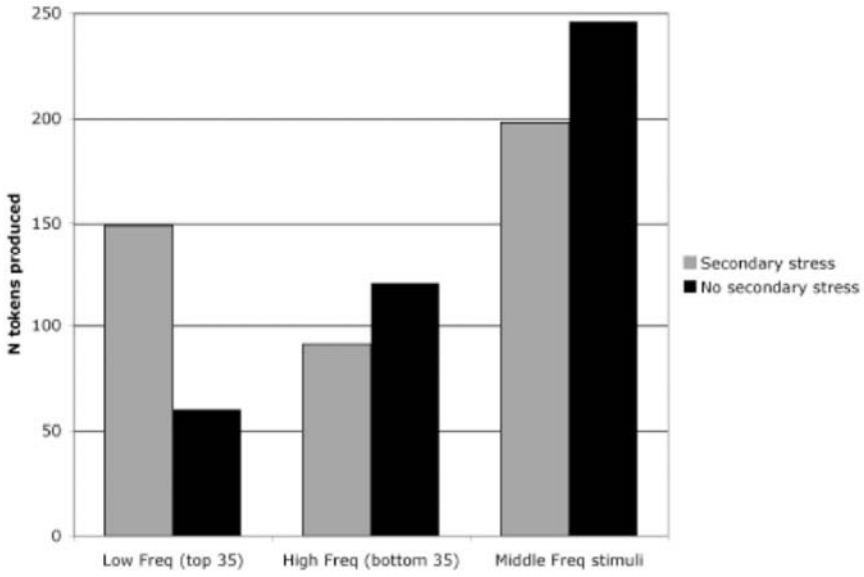
| STEM | | DERIVED FORM | |
|---|---|---|---|
| n[jé]ve | 'snow' | n[e]vóso | 'snowy' |
| verg[wé]nza | 'shame' | verg[o]nzóso | 'embarrassing' |
| v[jé]jo | 'old' | v[je]jecíto | 'little old man' |
| p[wé]blo | 'town, people' | p[we]blíto | 'little village' |
| cal[jé]nte | 'hot' | cal[je]ntíto | 'warm/cozy' |
| | | cal[e]ntíto | |

(p. 512)

The more productive the suffix (by corpus measures), the more likely to keep the diphthong.

- Gouskova & Roon 2008: in Russian compounds, the constraint requiring each stem to bear a prominence is ranked low, but there's a higher-ranked version of the constraint for low-frequency stems, forcing a secondary stress:

Figure 1: Effect of frequency on secondary stress realization




(p. 56)

## 8. Lexical information becomes available in real time

- How does that affect phonology that needs the lexical information? Student presentations of Wagner 2012.

**References**

Alcantara, Jonathan B. 1998. The Architecture of the English Lexicon. Cornell University.

Bermúdez-Otero, Ricardo. 2012. The architecture of grammar and the division of labour in exponence. In Jochen Trommer (ed.), *The morphology and phonology of exponence: the state of the art*, 8–83. Oxford: Oxford University Press.

Bybee, Joan L. & Dan I. Slobin. 1982. Rules and Schemas in the Development and Use of the English past Tense. *Language* 58(2). 265–289.

Carlson, Matthew T & Chip Gerfen. 2011. Productivity is the key: morphophonology and the riddle of alternating diphthongs in Spanish. *Language* 87(3). 510–538.

Chung, Sandra. 1983. Transderivational relationships in Chamorro phonology. *Language* 59. 35–66.

Coetzee, Andries W. 2008. Phonological variation and lexical frequency. *Porceedings of NELS 38*.

Collie, Sarah. 2008. English stress preservation: The case for "fake cyclicity." *English Language and Linguistics* 12(03). 505–532.

Gouskova, Maria & Kevin Roon. 2008. Interface constraints and frequency in Russian compound stress. In Jodi Reich, Maria Babyonyshev & Darya Kavitskaya (eds.), *Proceedings of FASL 17*. Ann Arbor, MI: Michigan Slavic Publications.

Hammond, Michael. 1999. Lexical frequency and rhythm. In Mike Darnell (ed.), *Functionalism and Formalism in Linguistics, Volume I: General Papers*, 329–358. John Benjamins.

Kirby, S. 2001. Spontaneous evolution of linguistic structure-an iterated learning model of the emergence of regularity and irregularity. *IEEE Transactions on Evolutionary Computation* 5(2). 102–110. doi:10.1109/4235.918430.

Löfstedt, Ingvar. 2010. Phonetic Effects in Swedish Phonology: Allomorphy and Paradigms. UCLA Ph.D. Dissertation.

Myers, James. 2005. Frequency effects and Optimality Theory. Handout. ROA #810-0306.

Ng, E-Ching. 2010. Reduction, frequency and morphology in Singaporean English prosody. Manuscript. Yale University, ms. ROA #1102.

Oxford English Corpus. Oxford English Corpus: Facts about the language. http://oxforddictionaries.com/words/the-oec-facts-about-the-language.

Pierrehumbert, Janet. 2001. Exemplar dynamics: word frequency, lenition, and contrast. In Joan Bybee & Paul Hopper (eds.), *Frequency effects and the emergence of linguistic structure*, 137–157. Amsterdam: John Benjamins.

Pinker, Steven. 2000. *Words and rules: the ingredients of language*. HarperCollins.

Wagner, Michael. 2012. Locality in phonology and production planning. In A McKillen & J Loughran (eds.), *Proceedings of the Montreal-Ottawa-Toronto (MOT) Phonology Workshop 2011. Phonology in the 21st Century: In honour of Glyne Piggott. McGill Working Papers in Linguistics 22(1).*

Zhang, Jie, Yuwen Lai & Craig Sailor. 2011. Modeling Taiwanese speakers' knowledge of tone sandhi in reduplication. *Lingua* 121. 181–206.