Bien & al. 2005, *PNAS*
Frequency effects in compound production

(1) What predicts how fast you can say *handbag*? Some candidates...
- left constituent family size: how many compounds start with *hand*?
- right constituent family size: how many compounds end with *bag*?
- left positional frequency: summed frequency of all compounds starting with *hand*
- right " " "
- left positional entropy: how evenly distributed are the token frequencies of compounds that start with *hand*?
- right " " "
- complement frequency: summed frequency of all <u>other</u> complex words containing *hand*
- derivational entropy: how even distributed are the frequencies of all the complex words that contain *hand*?
- lemma frequency: summed frequency of *hand*, *hands* (and any other inflected form)
- compound frequency: lemma frequency of compound (*handbag*, *handbags*)

(2) Experiments (Dutch)—comparisons
- Exp. 1: high vs. low head noun frequency: *luchtbrug* 'airlift' & *luchtbuks* 'airgun
- Exp. 2: high vs. low modifier noun frequency
- Exp. 3: both constituents high vs. low frequency
- Exp. 4: high vs. low compound frequency

(3) Experiments—method
- Learn to associate each member of a pair of compounds with a different position on the screen by hearing them over headphones and seeing a loudspeaker icon at the position
  - e.g., in Exp. 1, *luchtbuks* and *broodkruim*: both have low-frequency second member; "minimal phonological overlap, no obvious semantic relation, and [...] similar compound frequencies" (p. 17877)
- A couple of practice trials where you have to click on the correct loudspeaker icon.
- Test phase: icon appears and you have to say the compound; computer records response time
  - interspersed with distractor task: digit naming

(4) Experiments—results

Table 1. Mean latencies for Exps. 1–4

Exp.	Frequency	Mean, ms (%)	LH, ms	HL, ms
1	High	457 ± 111 (3)	437	476
	Low	471 ± 116 (2)	458	482
2	High	443 ± 118 (5)	439	447
	Low	468 ± 129 (5)	487	450
3	High	414 ± 105 (6)	405	424
	Low	441 ± 115 (5)	445	437
4	High	442 ± 108 (4)	430	454
	Low	434 ± 104 (4)	433	435

Values are for the main effect of frequency ± standard deviation (with error percentages in parentheses) and for the block orders low–high (LH) and high–low (HL).

(p. 17878)

- Exp. 1: higher right-const. frequency → faster
  - This is a little surprising: means you don't just get started on uttering the left constituent and worry about the second const. when you come to it.
  - Either the right const. has to get activated before you can start speaking...
  - ...or the resting activation of the whole compound depends on right-const frequency (let's think about whether that's plausible)
- Exp. 2: higher left-const. frequency → faster
  - Suggests synthetic access (at least sometimes)
- Exp. 3: both constituents more frequent → faster
  - even though freq. of whole compound matched within pairs
- Exp. 4: higher compound frequency doesn't make responses faster! (not fully significant though)

(5) What about other measures of productivity?

- Giant stepwise regression analysis.
- What's with "plosive"? The idea was that initial consonant type could affect the equipment's ability to pick up the response right away.
- This model does significantly better than one that has just the nonfrequency predictors (neighborhood density and consonant type) plus left and right const. cumulative root frequencies
  - interpreted as: there are effects here that a strict decomposition model doesn't capture

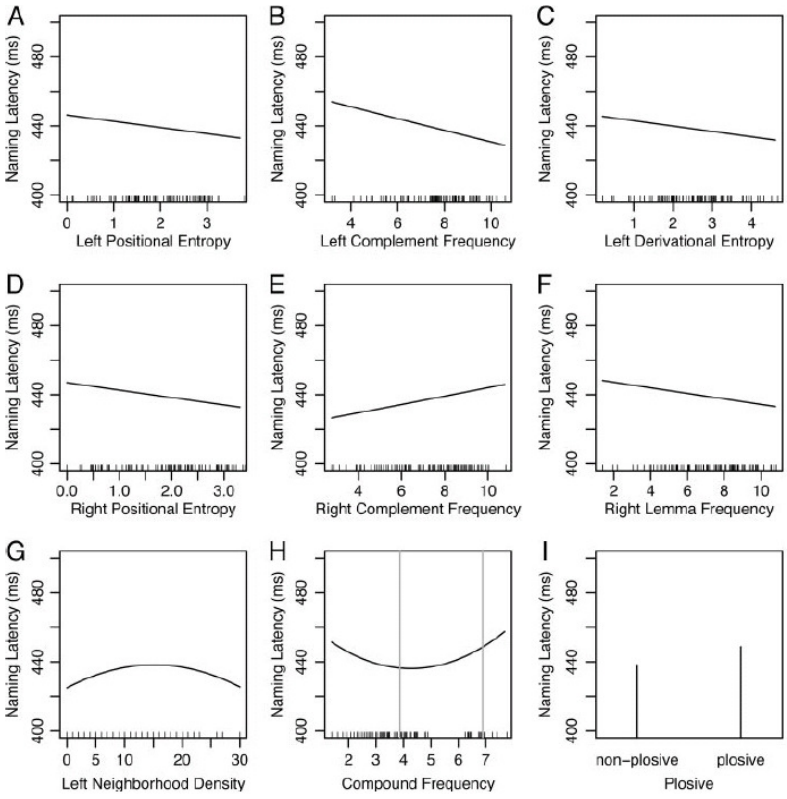

Fig. 1. Partial effects of the predictors in the multilevel covariance analysis of the data of Exps. 1–4. The left vertical axis shows the effect in log units; the right axis shows the effect in milliseconds. Values pertain to words that do not begin with a plosive and are adjusted for the effects of the other covariates at their median value. (p. 17880)

(6) Summary/discussion

- Frequency effects on production from compound's constituents (and not from whole compound)
  - not just full listing
- Frequency of second constituent matters too
  - "Speakers apparently plan the articulation of the first constituent with an eye on what is to be produced next" (p. 17881)
  - or at least, speakers don't start implementing production of the first constituent until access of the second succeeds (whether or not the *way* the first const. gets produced is affected)
- Contextual frequency measures: why should the number of compounds that begin with *hand* matter, as opposed to just the type or token frequency of any words containing *hand*?
  - Speculate that maybe this is because *hand* is pronounced differently when it's the modifier noun in a compound than elsewhere
- Positional entropy effects
  - Interpretation unclear.
- Inhibitory effect of right complement frequency
  - Could be problematic for right constituent to be getting activated while you're still trying to plan the left constituent.