Class 3 & 4 (Jan. 10 & 12): Models of lexical access in comprehension/perception

1 Overview

- Last time we looked at the lexical-access part of a couple of models of speech production.
- This time we'll do the same for listening or reading.
  - A lot of experimental methods present words visually, so we need to look a bit at reading regardless of whether we're interested in it per se.
- As before, the goal is not to understand the entire model—just the lexical access part.
- We also won't get too deeply into the debates about models' predictions and how well they line up with experimental findings—the goal is just to get the lay of the land before we start seeing how morphology fits it.

- Models to cover
  - Logogen
  - Cohort
  - Bin
  - Neighborhood activation
  - TRACE
  - Shortlist
  - One reading model: Coltheart & al.'s DRC

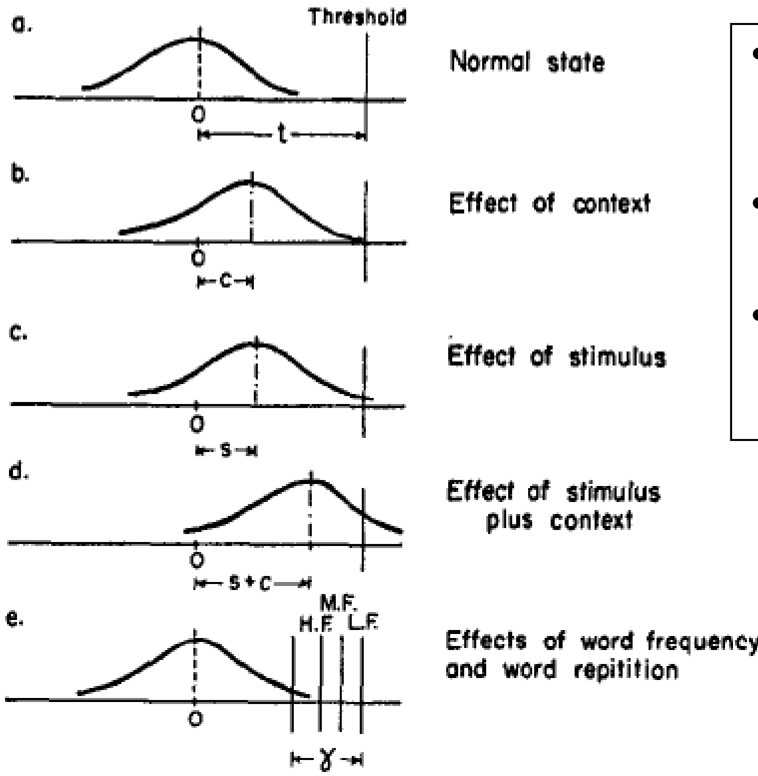
As we go, let's collect on the board dimensions that models can vary along.

2 Logogen model (Morton 1969)

(1) Basic model

- Each lexical unit (logogen) has phonological, syntactic, and semantic specifications
- Sensory input that matches the phonological specifications and context that matches the other specifications increases the logogen's activation
- First logogen to cross activation threshold wins and gets sent to the next stage of processing
  - losing logogens shouldn't affect later stages of processing
- Activation rapidly falls back to resting level

(2) Probability density functions for activation



- After a logogen has been selected, activation falls back to zero but threshold is lowered for a long time (→ priming)
- Frequency effects are therefore embodied not in resting activation but in frequently-lowered threshold.
- HF=high frequency, MF = medium frequency, LF=low frequency

(Morton 1969, p. 167)

(3) Later versions
- Might need separate logogens for speech input, reading input, and output
- This is to explain short duration of cross-modal repetition priming, compared to long duration of cis-modal priming.

3 Cohort model (Marslen-Wilson 1987)

(4) Three stages
- Access: map speech signal to lexical representation
- Selection: pick the best match
- Integration: use the lexical item for higher-level processing

Proposes that access is strictly bottom-up.
- only words consistent with the sensory input so far are in the cohort of contenders

Top-down factors influence selection from within the cohort
- → context can't activate a word that is inconsistent with sensory input

(5) Recognition point
- (English) You hear [t] → activate *tree*, *title*, *trespass*, *took*, *treadle*....
  - Now [tɹ] → eliminate *title, took*
  - Now [tɹɛ] → eliminate *tree*
  - Now [tɹɛs...] → eliminate *treadle*. Could still be *trespass*, *tress*, *trestle*
  - Now [tɹɛsp...] → the word can only be *trespass* (or a suffixed form of it)
- At any point, if context ruled out other competitors, we could have settled on *trespass* earlier.
- E.g., if you knew you were expecting a verb, [tɹɛs...] should be enough.[1]
- But at no point could you activate *invade* just because you were expecting it.

[1] OED gives also *tressilate*, but it's rare.

(6) Evidence from phoneme-monitoring task

Marslen-Wilson 1984 (didn't read), as cited by M-W 1987.

- "press the button as soon as you hear a *p*"
    - The idea is that you could press the button even before you've heard a *p*, once you know that the word is going to contain one
- correlation between response latency and uniqueness point in cohort
- correlation between response latency and recognition point (85% accuracy with subjects 85% confident) in gating task

(7) Evidence from lexical decision task

- "press the green button if it's a real word, red if it's not" (e.g.)
    - You hear [t] → could be *tree*, *title*, *trespass*, *took*, *treadle*....
    - Now [tɹ] → could still be *tree*, *trespass*, *treadle...*
    - Now [tɹɛ] → could still be *trespass, treadle, trend*, *trench...*
    - Now [tɹɛn...] → could still be *trend*, *trench...*
    - Now [tɹɛnk...][2] → can't be a word
- call /k/ in this example the "critical phoneme"

- Result: decision time was consistently about 450 ms. from the point just before the critical phoneme (/k/ above).
- Didn't matter if critical phoneme was in beginning, middle, or end of syllable
    - Suggests that units of access are phones, not syllables

(8) Competition

- In the cohort model, it's not just a race to cross an absolute activation threshold (cf. Morton's logogen model)
    - Instead, an item wins if it's the last one still activated, or perhaps if it outstrips all rivals by some minimum amount

o In an absolute-activation race, how would lexical decision be modeled?
o Would we expect recognition-point effects?

(9) Parallelism

- Forster 1976 proposes a model (that we'll look at below) that checks lexical items one by one.

o Let's brainstorm ways this could work efficiently
    - How do we decide which word to start with, when to move on from it, where to move to, and when to stop?

- By contrast, in the cohort model all the current competitors are dealt with at once

Consider lexical decision on these two non-words (abstracting away from much phonetic reality!), *trenk* and *bipse*. Counts from Weide 1993.

later decision point, smaller cohort		earlier decision point, bigger cohort	
T	5354	B	9244
T R	1164	B IH1	415
T R EH1	121	B IH1 P	0[3]
T R EH1 NG	0[4]		

---

[2] M-W's example. In this case place assimilation of the nasal would narrow it down a lot sooner
[3] Again, actually there's one match, *bippus*
[4] Actually there are 2: *trenchard* and *trenkle*, but they appear to be people's last names.

- What does our version of serial search predict about which will be faster?
- What does the cohort model predict about which will be faster?

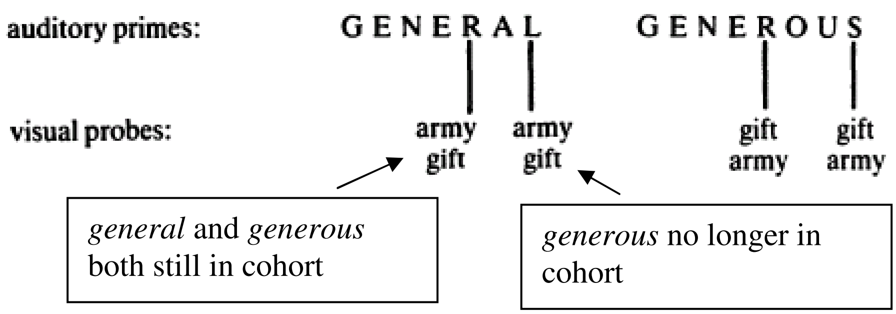
M-W reports that varying the "terminal cohort" size doesn't affect response time.

(10) Limited role for top-down information

- Under cohort model, top-down information can't prevent a word from being accessed
  - can only rule it out sooner
  - this might not even involve directly suppressing activation
  - if context only enhances activation, and winning is relative (just need to attain a certain share of the total activation), could just cause target to outstrip rivals faster

(11) Cross-modal priming, no context (Zwitserlood 1989)

- You hear *The next word is* <u>*general*</u> or *The next word is* <u>*generous*</u> (really in Dutch)
- At some point during the word, you see another word on screen (*army* or *gift*)
- You press a button to say whether what you saw was a real word
- If *general* is active, you should respond faster to *army*
- If *generous* is active, you should respond faster to *gift*



(Zwitserlood 1989, p. 31)

- means that activation is flowing to semantically related words before word recognition is finished

(12) Cross-modal priming, with context

Zwitserlood used three conditions—here the word you're going to end up hearing is *captain* (*kapitein*), and the competitor of interest is *capital* (*kapitaal*)..

- Carrier phrase: *The next word is* <u>*captain*</u> (word could be anything)
- Neutral context: *They mourned the loss of their* <u>*captain*</u> (still a lot of possibilities)
- Biasing context: *With dampened spirits the men stood around the grave. They mourned the loss of their* <u>*captain*</u> (more likely to be *captain* than *capital*)
- Control: *The player got the ball, and scored the winning* <u>*goal*</u> (neither *captain* nor *capital* should be especially activated)

You'll be asked to judge *money* or *ship* (call that the "probe")

(13) A bit more context

When in the word do you see the probe?

- Position zero: just before the word starts

Conducted a gating pre-test

- Play the sentence, but only the first 50 ms. of the last word
- Ask subject to name the word and rate their confidence in it
- Then play the first 100 ms of the word
- Then the first 150 ms, etc.

- Probe position 1: average "isolating point"—earliest point at which subjects produced the right word, without changing their minds later—in the Biasing context
  - 130 ms. into the word, on average
- Probe position 2: average isolating point in Neutral context
  - 199 ms. into the word, on average
  - In Carrier phrase condition, subjects in pre-test produced on average 7.5 alternatives at Position 1 and 6.5 at Position 2
- Probe position 3: mean isolation point in Carrier phrase condition
  - 278 ms. in, on average
  - majority of pre-test subjects producing right word, but still some alternatives (3.2 on average)
- Probe position 4: mean point at which subjects produced the correct word (in Carrier phrase condition), with confidence rating of 9 or 10 out of 10, and didn't change mind later.
  - 410 ms. in, on average

"zero condition": you're shown the probe before you start hearing any of *captain* or *capital*.

(14) Zwitserlood results

- Context doesn't matter if you haven't heard any of the word yet

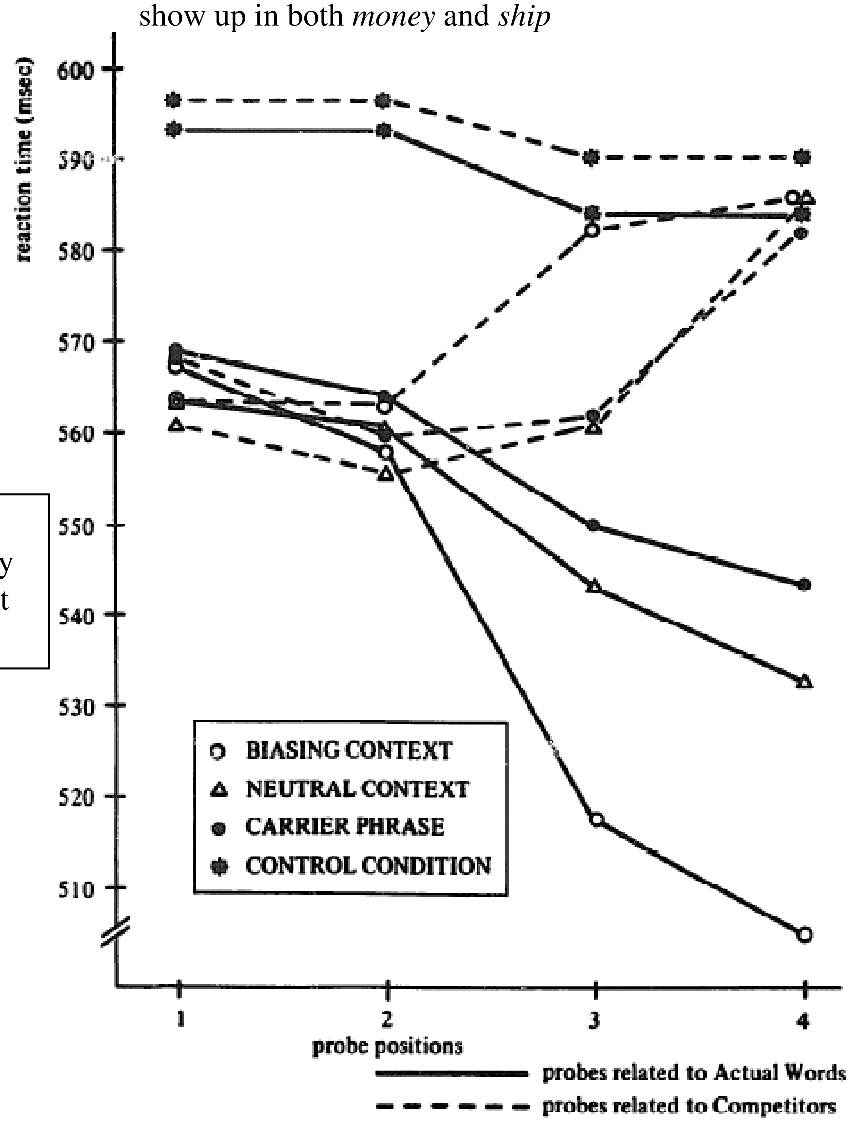
	Probe related to:	Zero condition Biasing Context	Zero condition Neutral Context	Control condition
ship	Actual Word	593	596	594
money	Competitor	592	590	596

(p. 45)

other conditions on next page

- Dashed line is words like *money*; solid line is words like *ship*
- Lower number means subjects were faster to access the probe and judge it a real word
- Compare Carrier to Control at Positions 1 and 2
  - Early on, responses to both *money* and *ship* are facilitated by hearing *kap*...
- Follow the curve for Carrier
  - At Position 3, *ship* is facilitated more than *money*
  - At Position 4, *ship* is even more facilitated, but *money* similar to Control
- Follow the curve for Neutral context
  - similar to Carrier
- Follow the curve for Biasing context
  - Despite top-down bias for *captain*, *money* is still facilitated at Positions 1 and 2
    - → context can't prevent *capital* from being activated
  - Despite top-down bias for *captain*, no additional facilitation of *ship* at Positions 1 and 2
    - → captain not activated until you start hearing it, and still has to compete with *capital*
  - Compare to Carrier and Neutral at Position 3: *money* no longer facilitated, *ship* very much facilitated
    - → context can eventually decide the race sooner
    - could be because context eventually suppresses *money*
    - could be because context eventually boosts *ship*
    - if winning means acquiring a big-enough share of total activation, either effect would show up in both *money* and *ship*

See table above: they all start out right here



(Zwitserlood p. 46)

(15) Frequency effects?

- No place for them in original Cohort model
- M-W 1987 does find that, even measuring lexical-decision times from recognition point (not from beginning of word), higher-frequency words are recognized more quickly
- In work-in-progress cited by M-W 1987, evidence that frequency relative to closest competitor matters

→ M-W 1987 proposes to incorporate differences in resting activation into model

(16) Robustness to errors?

- Shadowing task: listen to recorded speech and repeat along as fast as you can
- Errors in the recorded speech tend to be repaired, and not even noticed (Marslen-Wilson & Welsh 1978, cites also Cole 1973 for direct error detection)
- In daily life, even if we notice a speech error it's usually not fatal to comprehension.
- How do words inconsistent with the speech signal get into the cohort?

→ M-W 1987 proposes that input is not as abstract as a string of phonemes

  - He doesn't quite say this, but if input is a distribution over phonemes (or other categories), and if being in the cohort is not a binary property, then mismatches will be penalized at first but perhaps boosted later by context (especially if the error produces a non-word).

(17) Ability to handle unexpected words

- *John buried the <u>guitar</u>*
  - M-W 1987 reports that response to *guitar* is 27 ms. slower than normal
- *John drank the <u>guitar</u>*
  - response 49 ms. slower than normal
- *John slept the <u>guitar</u>*
  - response maybe about 25 ms. slower still than that, but doesn't say exactly
- → Unexpectedness slows us down, but not that much, and rarely leads to perception errors

- Follows from bottom-up priority in model
- M-W concludes that context can't exclude items from cohort (just as it can't add items to it)
- Tentatively proposes that lexical selection balances activation levels that reflect fit to the sensory input (and starting activation), and goodness of fit to context.

4 Bin model (Forster 1976)

We can see how close our guesses were...

- Forster is particularly concerned with explaining lexical-decision behavior
  - How do we ensure that *destair* results in a "no" response, rather than activating *despair* (the closest match under some measures) and saying "yes"?
  - Nonwords similar to real words do make the decision slower (various refs in Forster), so the real words are getting activated to some degree—they just shouldn't win.

- One master "file", that is indexed three ways (for reading, listening, and speaking)

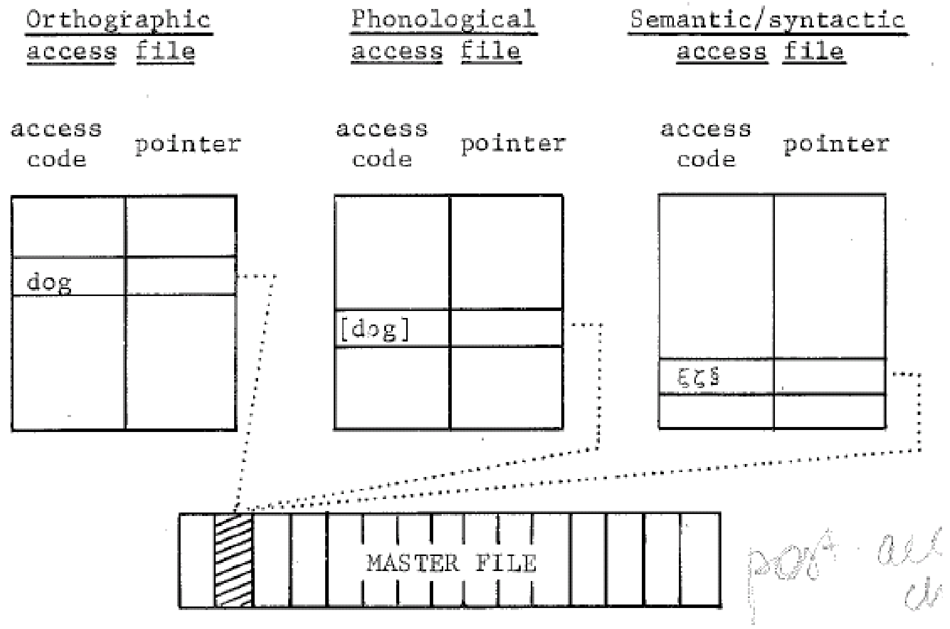


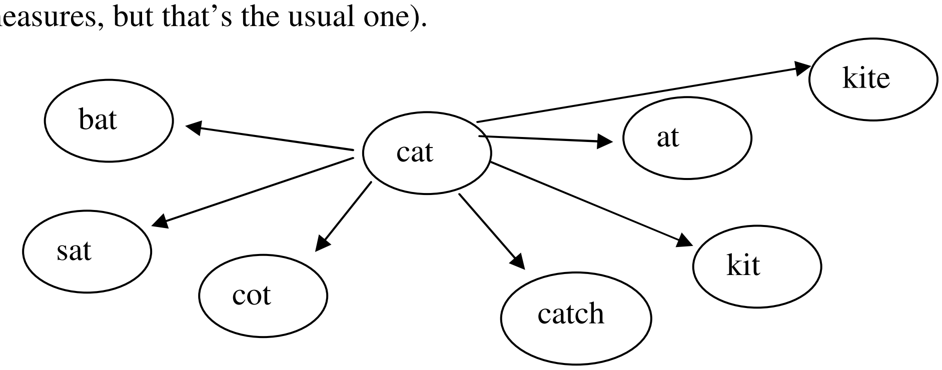
Figure 4. Organization of peripheral access files and master file. (Forster 1976, p. 268)

- Access file is grouped into bins
- Each bin is organized by frequency (most frequent at the "top"), or maybe age of acquisition
- To access, find the right bin, start at the top, and keep going till you find an item that exceeds some similarity threshold to the access code
- Look up the matching item in the master file to check it.
  - if it's not a perfect match, keep looking
  - What happens when you read a word that you yourself always misspell?
  - eventually, you can give up
- How are the bins organized?
  - Forster gives some speculations, e.g. maybe one bin for each combination of first and last letter (for reading)

5 Neighborhood activation model (Luce & Pisoni 1998)

(18) Neighborhood

A word's <u>neighborhood</u> is the words that differ from it by one phone (we could have other distance measures, but that's the usual one).



(19) The basic idea

- If the cohort of competitors is actually the neighborhood, then recognition should be slowed down if the neighbors are numerous and/or frequent.
- You can separately manipulate neighborhood size, average neighbor frequency, and target word frequency.

(20) Experiment

- Listen to a C_1VC_2 word (with various amounts of noise) over headphones
- Identify the word by typing it.
- Luce & Pisoni first use the overall data to construct a confusion matrix for C_1, V, C_2.
- Model probability of correct identification as:

$$p(identify(stimulus)) = \frac{WP(stimulus)}{WP(stimulus) + \sum_{j=1}^{m} WP(N_j)}$$, where each N is a neighbor of the stimulus

and $WP(word_k) = freq(stimulus) \times \prod_{i=1}^{n} p(identify(word_k Phoneme_i) \mid play(stimulusPhoneme_i))$

- In other words...
  - For each word, multiply the probabilities of hearing that word's C_1, V, and C_2 given that the target word's C_1, V, and C_2 were played.
  - Multiply that by the word's frequency, and call that the word's WP.[5]
  - Divide the target word's WP by the sum of all the WPs in the neighborhood (including its own)
- Result: the equation above is much better correlated with correct identification rates than is log word frequency (except in the noisiest stimuli, where both correlations are equally so-so).
  - Doesn't look like they tried correlating just WP (frequency and phonetic identifiability) to the results.

(21) Model

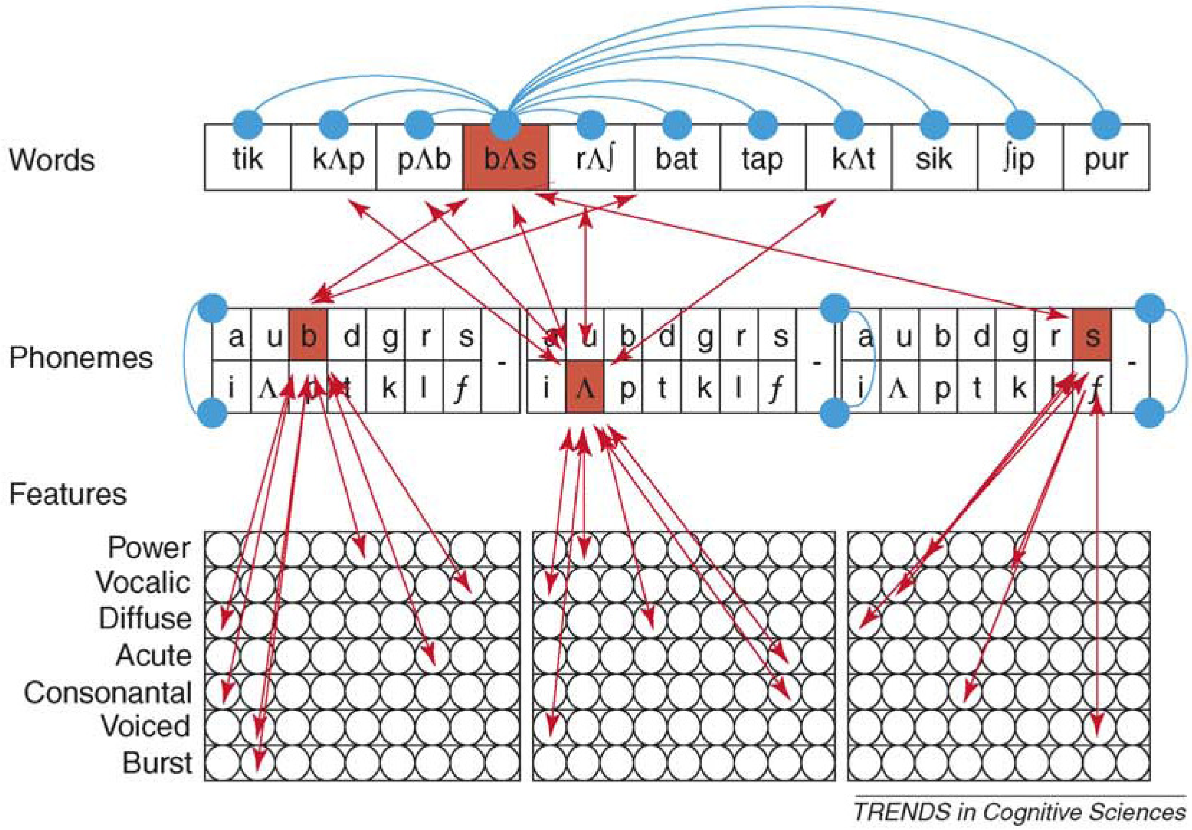
- Similar to Morton's logogen model: each lexical item is a "word decision unit"
  - unit is activated if its form specifications are close enough to the sensory information coming in (→ set of units activated is approximately the target's neighborhood)
  - unit then monitors higher-level information (e.g., sentence context; also frequency)
  - if unit's share of activation exceeds some threshold, then the item is activated and all the lexical information made available to working memory
- o Let's again discuss differences between this model and the others we've seen.

[5] Luce & Pisoni set it up a bit differently, but I think this is totally equivalent.

6 TRACE model (McClelland & Elman 1986)

Discussion here based mainly on McClelland, Mirman, & Holt 2006, because it's more recent.

(22) The model

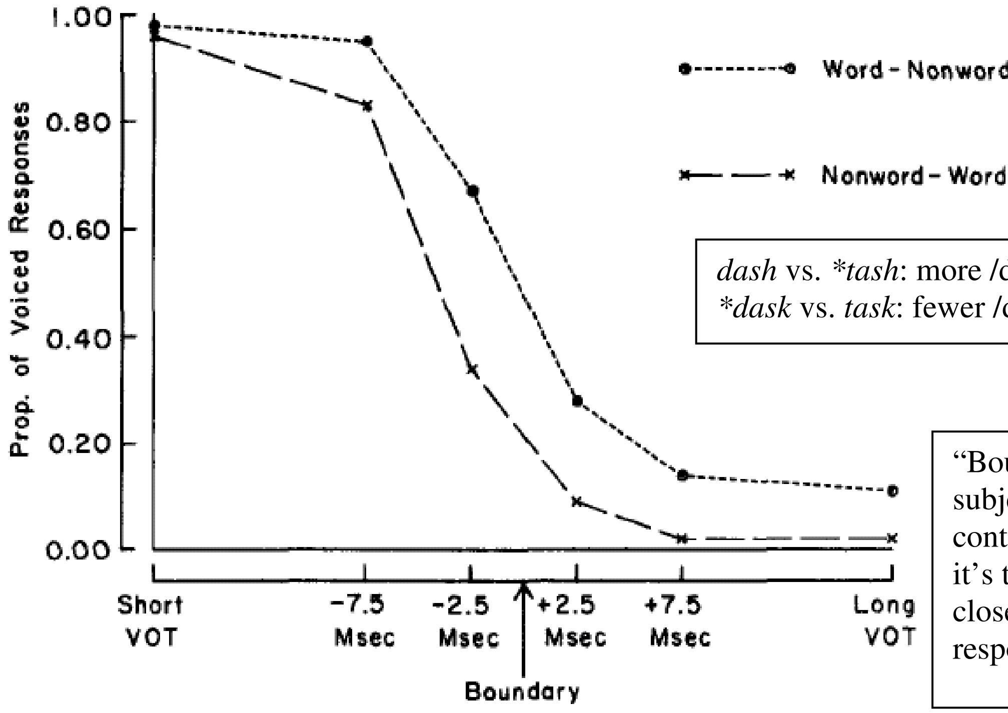


(McClelland & al. 2006, p. 365)

- Connections with arrows are excitatory
- Connections with circles are inhibitory (competing units)
- Connections are all bidirectional → interaction (recall Dell vs. Levelt for production). So let's look at the evidence for that...

(23) Top-down effects?

- Ganong 1980: speech perception is influenced by lexicon
  - Construct an ambiguous [d]/[t] sound; call it [D].
  - [Dæʃ] tends to get perceived as *dash*
  - [Dæsk] tense to get perceives as *task*



dash vs. **tash*: more /d/ responses
**dask* vs. *task*: fewer /d/ responses

"Boundary": for each subject and each continuum (t-d or k-g), it's the VOT with closest to 50% voiced responses.

Figure 3. Results of Experiment 1. (Phonetic categorizations pooled as described in the text.)
(Ganong 1980, p. 115)

- True interactivity, or bottom-up followed by later lexical influences?
- Looking for evidence about whether lexical effects happen <u>early</u>.

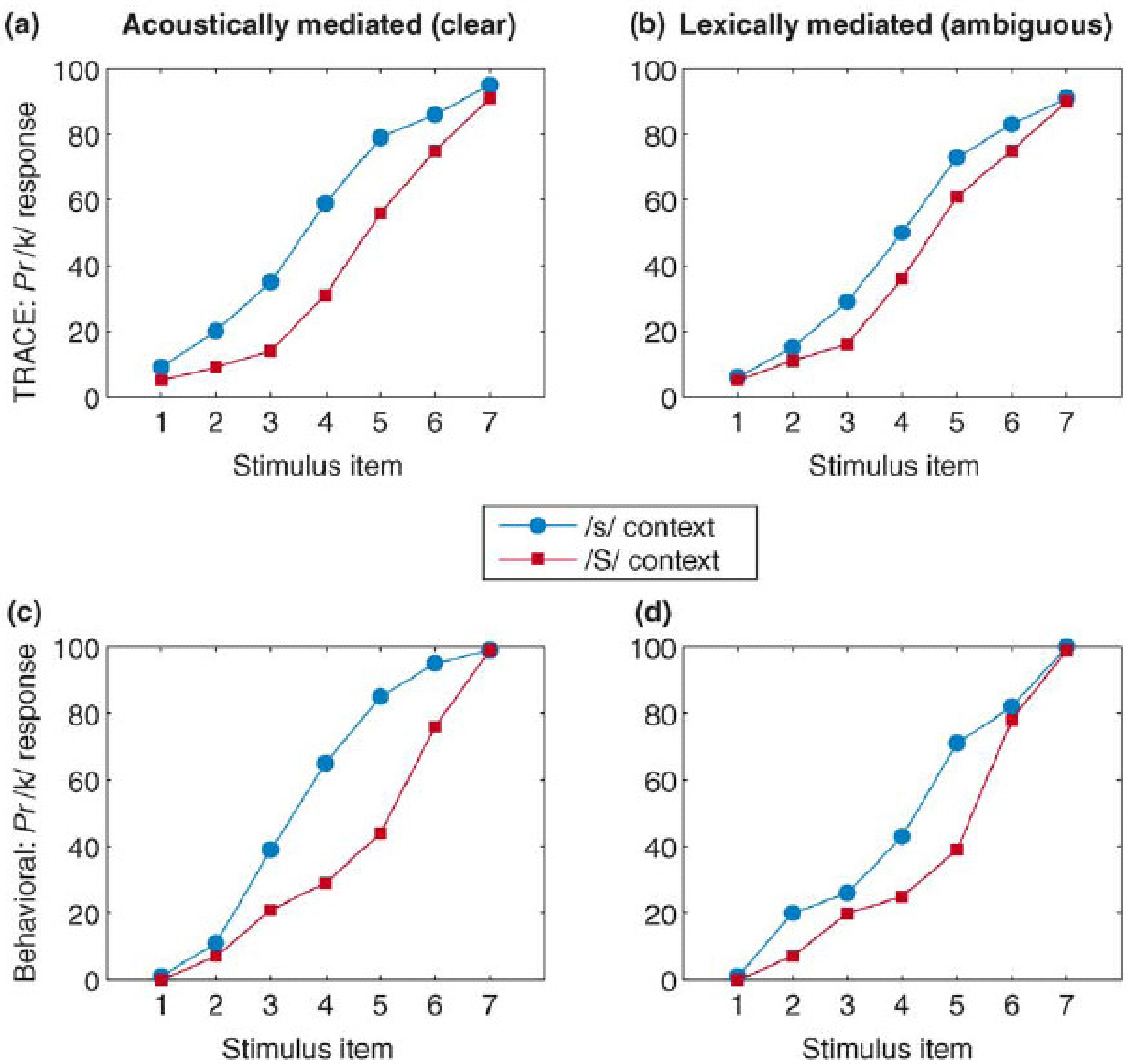
(24) Evidence 1: compensation for local auditory context

Elman & McClelland 1988, as reported in McClelland & al. 2006

- Shown elsewhere (Mann & Repp 1980): ambiguous [t]/[k] (call it [T]) sound heard as k/s__ and as t/ʃ__
- Produce ambiguous [s]/[ʃ] sound; call it S.
- In *Christma[S]* it'll be heard as /s/; in *fooli[S]* it'll be heard as /ʃ/.
- What about a following [T]?

(a) & (b): TRACE model

(c) & (d): humans



TRENDS in Cognitive Sciences

More /k/ heard after actual [s] than after actual [ʃ]

More /k/ heard after ambiguous [S] when lexical context requires [s] than when lexical context requires [ʃ]

McClelland & al. 2006, p. 367

(25) Other arguments: selective adaptation and retuning phonemic categories

I won't go through these, but they similarly involve arguments that lexical effects happen at stages supposed to be pre-lexical, and that this requires interactivity.

See McQueen et al. 2006 for a rebuttal.

(26) Left-to-right stuff

- The information available to the listener unfolds over time.
- So, at first the words getting the most activation will be those that share the beginning of the target word.
- In that sense it's similar to the Cohort model
- o What are some differences from Cohort?

(27) Let's play with an implementation on screen and see if we can make sense of it

Strauss, Harris, & Magnuson 2007's jTRACE, http://maglab.psy.uconn.edu/jtrace/

7 Shortlist (Dennis Norris 1994)

See also Norris & McQueen 2008 for Shortlist B, which involves calculating Bayesian probabilities rather than passing activation.

(28) Has elements of both Cohort and TRACE

- First derive a shortlist of items consistent with sensory input (bottom-up first, like Cohort)
- Then let activation flow among those items in a TRACE-like way, but without the phoneme nodes
  - → Unlike TRACE, no downward activation flow from lexical to phonological nodes

(29) Explanation for, e.g., lexical influences on phoneme monitoring?

Similar to Cutler & Norris 1979's race model (which Norris says Shortlist is sort of an implementation of: pp. 207-208).
See also the Merge model (McQueen et al. 2006)

- Phoneme monitoring ("press the button when you hear a *p*") involves two racing routes
  - wait for the phone node to get activated by sensory information
  - look for the phone in an activated lexical representation
- So, you can spot the phone early if you've activated the right lexical item, without sending activation down to the phone itself

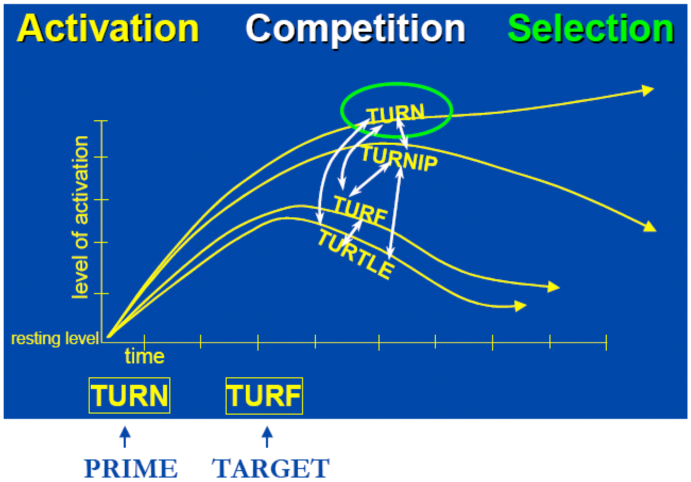
8 Miscellaneous concluding thoughts on models of lexical access in speech

(30) Syllables

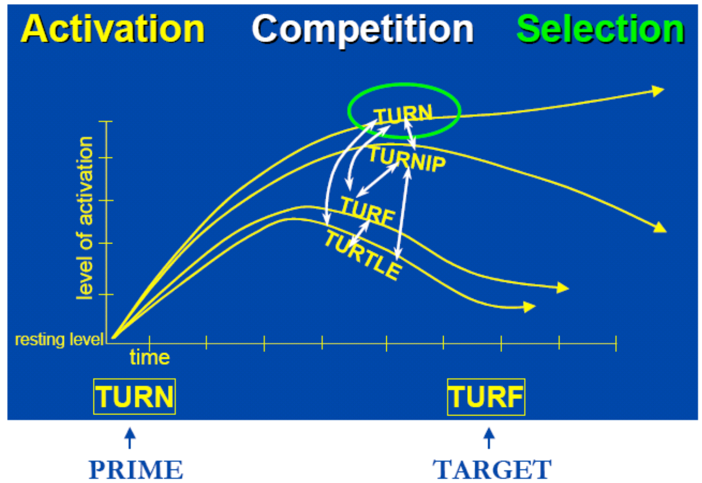
- Some models have them, some don't.
- How much should a morpheme's recognizability suffer under resyllabification?
  - *artist* + *ic* → *ar.tis.tic* (or maybe *ar.ti.stic*): harder to recognize *artist* and *–ic*?
  - see Raffelsiefen 2004 for argument that V-initial and C-initial suffixes in English appear to be associated with different lexical levels because they trigger different p-word structures
- Importance of syllables could be language-specific (e.g., Otake et al. 1993 on English stress vs. French syllables vs. Japanese moras)
- Or should we be thinking not in terms of syllables but in terms of all kinds of non-contrastive features?
  - some syllabification might be purely performed by the listener
  - but much has perceptible effects on segments (e.g., English aspiration)
  - so resyllabification is in the same boat as other phonological changes that a morpheme undergoes because of neighboring morphemes

(31) Timecourse of phonological priming

A picture by Liina Pylkkänen (homepages.nyu.edu/~mp108/Neural_Bases_of_Language_Fall2009.html) illustrates nicely:



TURN facilitates TURF, because TURF is still in the running, receiving activation from sensory input.

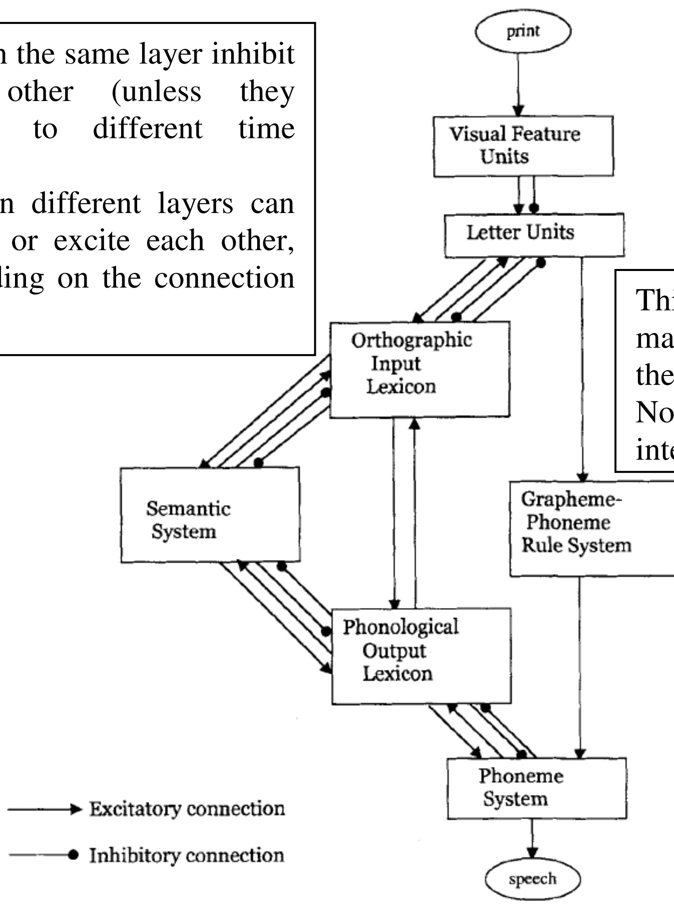


By now, TURF's activation has been suppressed by successful competitor TURN, so there's no priming, or even slowing compared to control if TURF's activation ends up lower than it was before.

9 Dual Route Cascaded (DRC) model for (alphabetic) reading (Coltheart et al. 2001)

(32) Basic architecture

- units in the same layer inhibit each other (unless they belong to different time slices)
- units in different layers can inhibit or excite each other, depending on the connection type



This bypass allows for sounding out new words, and can make real words get read aloud faster the more regularly they're spelled.
Not exactly a dual-route race model: the two routes interact.

As in Levelt's production model, homophones have a single phonological unit—and homographs have a single orthographic unit

(Coltheart & al. p. 214)

(33) Activation dynamics

- Similar to what we've seen in other spreading-activation models:

$$a_i(t+\Delta t) = a_i(t) - d_i(a_i(t)) + \varepsilon_i(t)*r$$ (similar to Coltheart & al.'s (1))

$a_i(t+\Delta t)$	$a_i(t)$	$d_i(a_i(t))$	$\varepsilon_i(t)*r$
activation at next timestep	activation now	decay	net incoming activation, times activation rate

- Except that incoming activation can be positive or negative
- If result is >1, corrected to 1; if <0, corrected to 0 (actually, that happens inside the $\varepsilon_i(t)$ function, but I think the difference is just algebra).
- Activation of letter-feature units is "clamped", but from then on it goes in both directions.

(34) Performance

- Does <u>really</u> well at reading CELEX aloud—only mistake is *czars* (pronounces it /kɑz/).
- 98% accurate and sounding out non-words too
  - most errors involve saying a real word instead ("lexical captures")
- Simulated various effects in human reading-aloud performance
  - e.g. frequency effect, interaction of regularity and frequency .... (I don't list them all because the tasks used to probe morphology rarely involve reading aloud)
- Simulated effects in lexical decision—I'll refer you to the paper for the results
  - faster for high-frequency
  - faster for "yes" answers than for "no"
  - neighborhood size (*N*): bigger *N* → faster "yes" when low frequency
  - *N* unimportant for "yes" when high frequency
  - bigger *N* → slower "no"
  - sounds out to real word (e.g., *trane*) → slower "no"; effect smaller when spelling is very different between fake word and homophonous real word

10 Wrapping up

- Monday is a holiday
- Wednesday we'll start with presentations of articles setting up general issues/dichotomies in lexical access of morphologically complex words
- Let's decides who's presenting what for the next couple of meetings (separate handout)

11 References

Cole, R. A. 1973. Listening for mispronunciations: a measure of what we hear during speech. *Perception and Psychophysics* 13. 153-156.

Coltheart, Max, Kathleen Rastle, Conrad Perry, Robyn Langdon & Johannes Ziegler. 2001. DRC, : A Dual Route Cascaded Model of Visual Word Recognition and Reading Aloud. *Psychological Review* 108(1). 204-256.

Cutler, A. & D. Norris. 1979. Monitoring sentence comprehension. In W.E. Cooper & E. C. T. Walker (eds.), *Sentence processing: psycholinguistic studies presented to Merrill Garrett*. Erlbaum.

Elman, Jeffrey L. & James L. McClelland. 1988. Cognitive penetration of the mechanisms of perception: Compensation for coarticulation of lexically restored phonemes. *Journal of Memory and Language* 27(2). 143-165. doi:10.1016/0749-596X(88)90071-X.

Forster, K. I. 1976. Accessing the mental lexicon. In , *New approaches to language mechanisms*. Amsterdam: North-Holland.

Ganong, William F. 1980. Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6(1). 110-125. doi:10.1037/0096-1523.6.1.110.

Luce, Paul A. & David B. Pisoni. 1998. Recognizing spoken words: the neighborhood activation model. *Ear and Hearing* 19(1). 1-36.

Mann, Virginia A. & Bruno H. Repp. 1980. Influence of vocalic context on perception of the [S]-[s] distinction. *Perception and Psychophysics* 28(3). 213-228.

Marslen-Wilson, William D. 1984. Function and process in spoken word-recognition. In H. Bouma & D. G. Bouwhuis (eds.), *Attention and Performance X: Control of language processes*. Hillsdale, NJ: Erlbaum.

Marslen-Wilson, William D. 1987. Functional parallelism in spoken word-recognition. *Cognition* 25(1-2). 71-102. doi:10.1016/0010-0277(87)90005-9.

Marslen-Wilson, William D. & Alan Welsh. 1978. Processing interactions and lexical access during word recognition in continuous speech. *Cognitive Psychology* 10(1). 29-63. doi:10.1016/0010-0285(78)90018-X.

McClelland, James L. & Jeffrey L. Elman. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18(1). 1-86. doi:10.1016/0010-0285(86)90015-0.

McClelland, James L., Daniel Mirman & Lori L. Holt. 2006. Are there interactive processes in speech perception? *Trends in Cognitive Sciences* 10(8). 363-369. doi:10.1016/j.tics.2006.06.007.

McQueen, James M., MARCS Auditory Laboratories, Anne Cutler & Dennis Norris. 2006. Are there really interactive processes in speech perception? http://handle.uws.edu.au:8081/1959.7/34000 (10 January, 2011).

Morton, John. 1969. Interaction of information in word recognition. *Psychological Review* 76(2). 165-178.

Norris, Dennis. 1994. Shortlist: a connectionist model of continuous speech recognition. *Cognition* 52(3). 189-234. doi:10.1016/0010-0277(94)90043-4.

Norris, Dennis & James McQueen. 2008. Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review* 115(2). 357-395.

Otake, T., G. Hatano, A. Cutler & J. Mehler. 1993. Mora or Syllable? Speech Segmentation in Japanese. *Journal of Memory and Language* 32(2). 258-278. doi:10.1006/jmla.1993.1014.

Raffelsiefen, Renate. 2004. Paradigm Uniformity effects versus boundary effects. In Laura J. Downing, T. Alan Hall, & Renate Raffelsiefen (eds.), *Paradigms in Phonological Theory*, 211-262. Oxford University Press.

Strauss, Ted J., Harlan D. Harris & James S. Magnuson. 2007. jTRACE: A reimplementation and extension of the TRACE model of speech perception and spoken word recognition. *Behavior Research Methods* 39(1). 19-30.

Weide, Robert. 1993. *Carnegie Mellon Pronouncing Dictionary [cmudict.0.2]*.

Zwitserlood, Pienie. 1989. The locus of the effects of sentential-semantic context in spoken-word processing. *Cognition* 32(1). 25-64. doi:10.1016/0010-0277(89)90013-9.