The 30th Annual Conference of the Cognitive Science Society

Computational Modeling of Spoken Language Processing: A hands-on tutorial

THE NEW SCHOOL

Haskins Laboratories

Computational Modeling of Spoken Language Processing: A hands-on tutorial

Ted Strauss

New School University

Dan Mirman

Jim Magnuson

University of Connecticut Department of Psychology

and

Haskins Laboratories

Plan

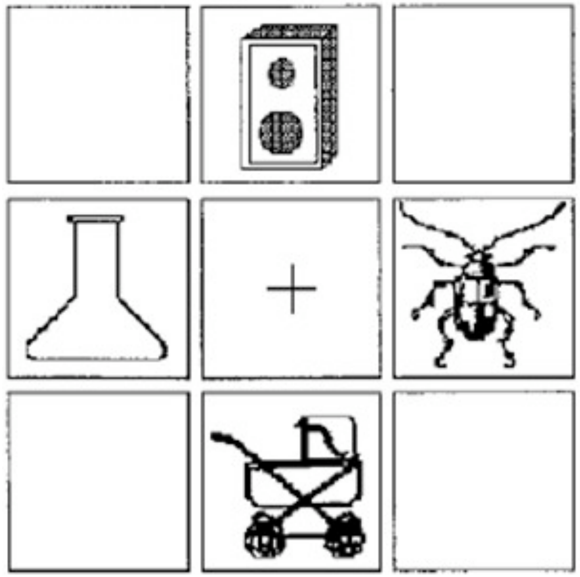
- Module 1: Introduction, About TRACE
- Module 2: Tour of jTRACE
- Module 3: Classic simulations
- Module 4: Scripting
- Module 5: Linking hypotheses
- Module 6: Lab time, Q&A, one-on-one

Module 3

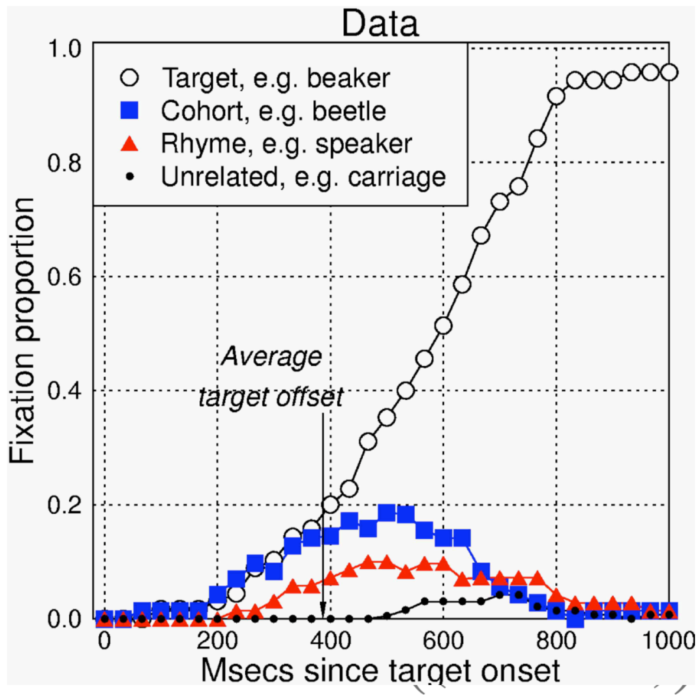
- Walk through simulations of classic effects
  - Demonstrate basic jTRACE simulation and analysis techniques
  - Effects at different levels to show off the breadth of TRACE's coverage
  - Complexity of simulations will gradually increase to show off jTRACE's capabilities

- Three cases:
  1. Time course of lexical activation and competition
  2. Lexical consequences of acoustic deviations
  3. Lexical effects on identification of ambiguous phonemes

1. Time Course of Lexical Activation and Competition



- Basic behavioral finding(s):
  - Early in word processing, words with similar onsets ("cohort", e.g., *beaker – beetle*) compete with the target
  - Later in word processing, words with similar offsets (rhyme, e.g., *beaker – speaker*) compete with the target
  - Cohort competition is stronger than rhyme competition



THE NEW SCHOOL

Simulation plan

1. Make critical target-cohort-rhyme triple
2. Run simulation with target as input
3. Analyze results: Linking hypothesis
   - Behavior: fixation proportion to 4 images on the screen (target, cohort, rhyme, unrelated) ≈ 4AFC
   - Model: convert lexical unit activations to forced choice probability among 4 lexical items using the Luce (1959) choice rule:

~ ratio of unit exponential activation relative to activation of other units

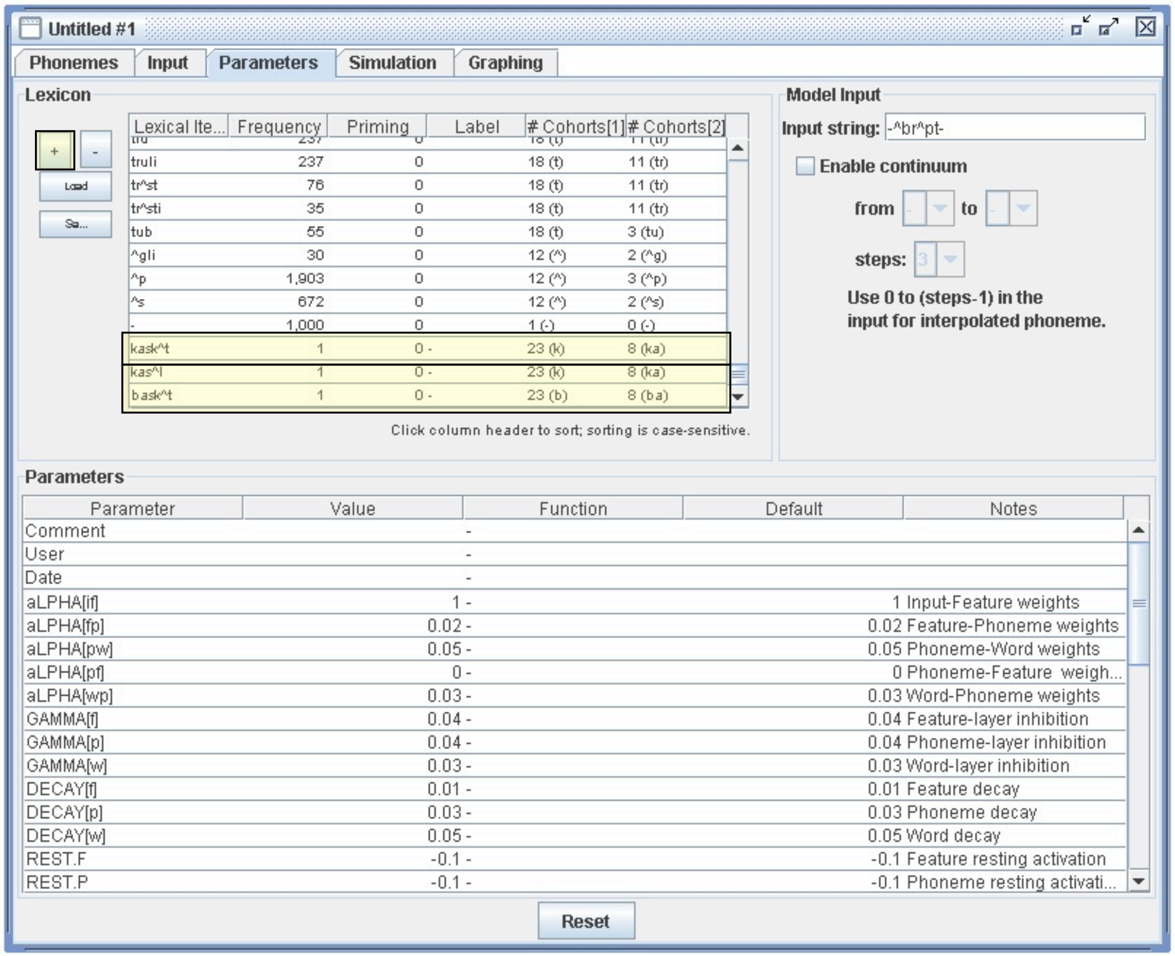
k = response competition

$$L_i = \frac{e^{ka_i}}{\sum_j e^{ka_j}}$$

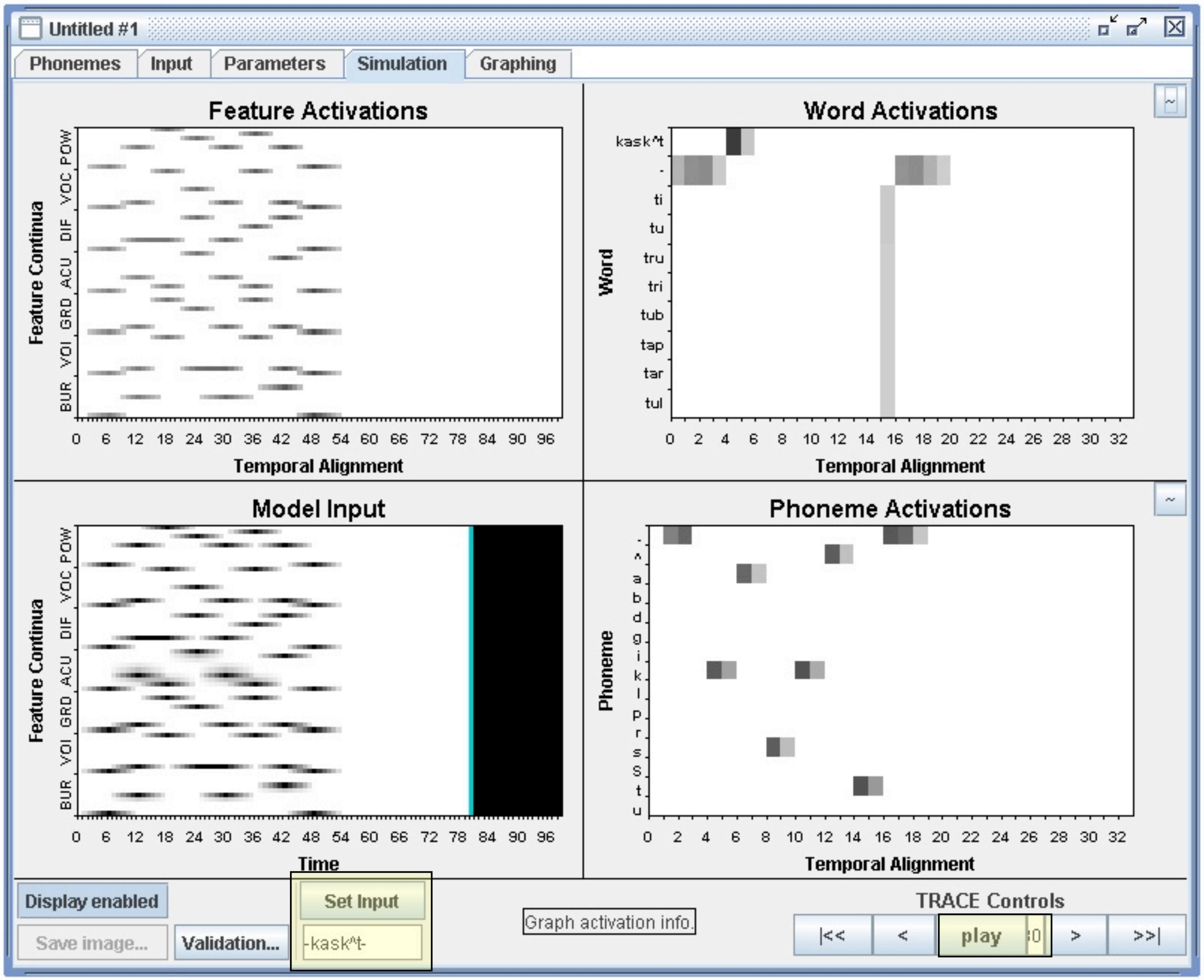
Simulation

1. Make a target-cohort-rhyme triple

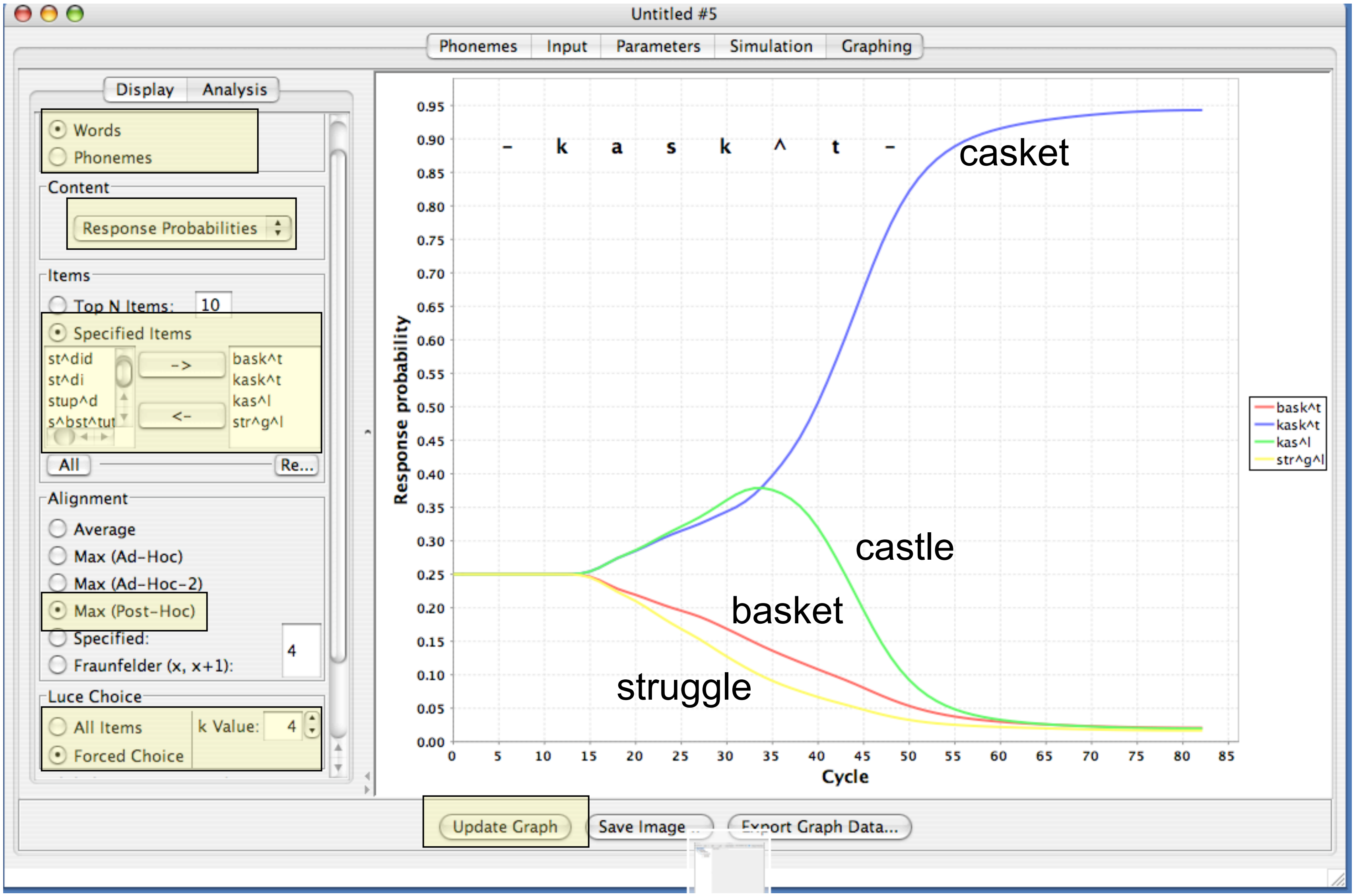
casket (kask^t), castle (kas^l), basket (bask^t)



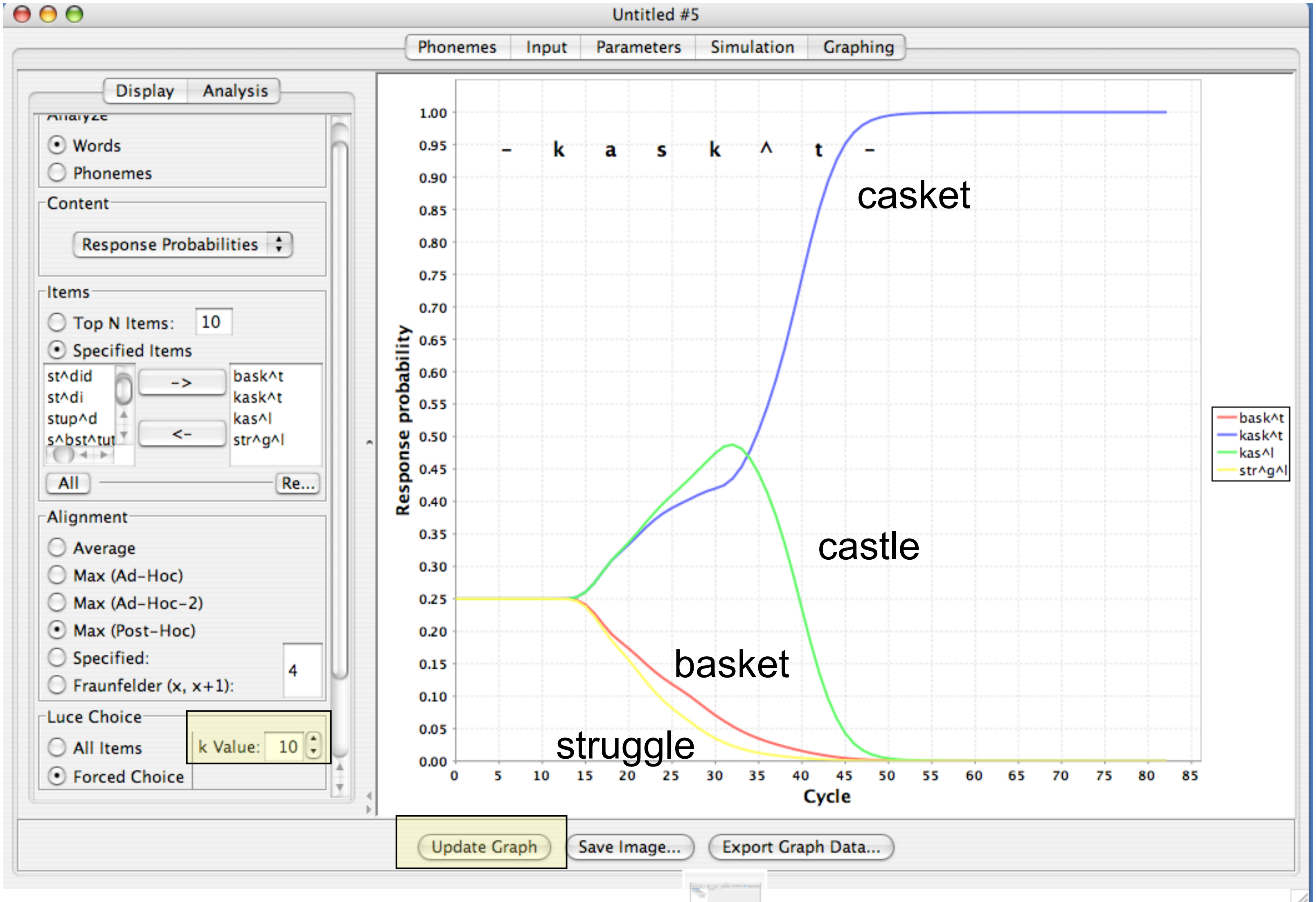
2. Run simulation



3. Analyze results



Enhanced Response Competition



Why does the model exhibit this pattern?

- Continuous mapping
  - Input matching any part of a word partially activates it
    - ➢Initial mismatch does not rule out a word
  - Thus, both cohorts and rhymes can become active
- Lateral inhibition
  - Rhyme competitors have to compete against already active words (target + cohort competitors), so they do not reach high activations

2. Lexical Consequences of Acoustic Deviations

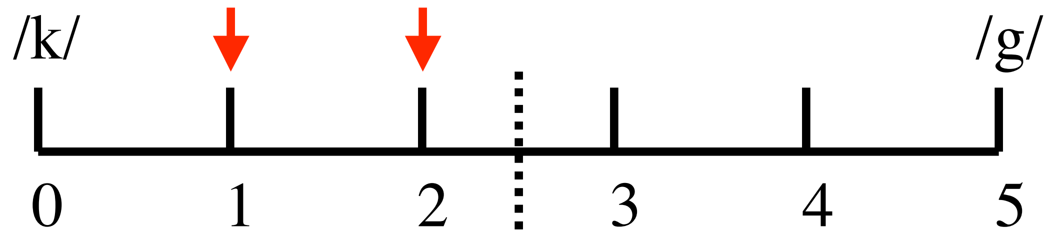
- Research question: is output of phonological processing (input to lexical processing) discrete or graded?
  - Traditional categorical perception claims that it is discrete
- Behavioral test (e.g., Andruski et al., 1994)
  - Make a /k/-/g/ continuum
    - ➢Critical test items: speech sounds (/k*/) that are identified as "k", but are a bit closer to /g/
  - Test activation of /k/-words when input contains good /k/ vs. deviant /k*/
    - ➢Word activation assessed by priming: /kar/ primes "truck" more than /k*ar/ does
    - ➢This difference disappears at longer SOA

Simulation Plan

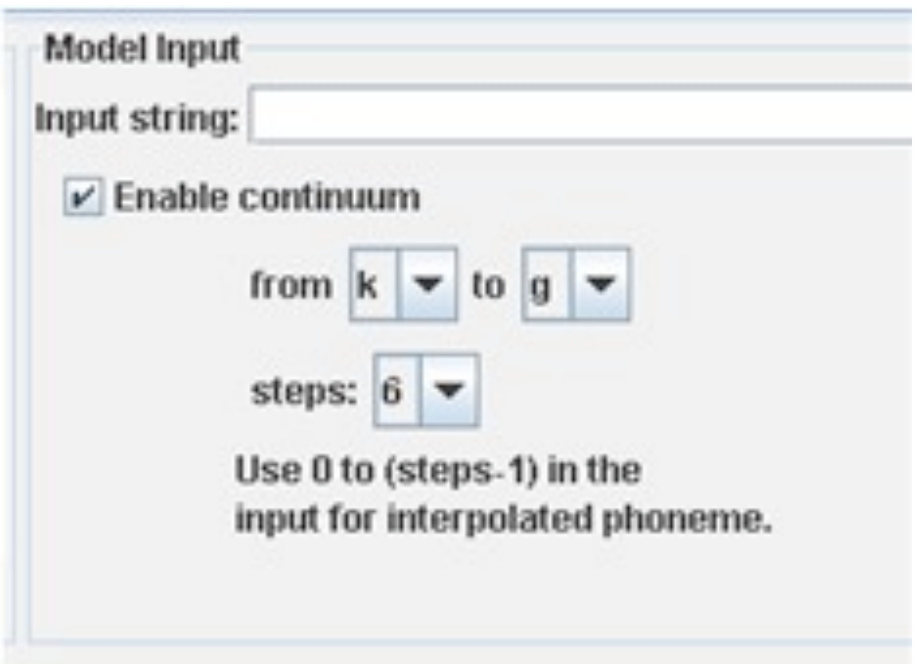
1. Create a /k/-/g/ continuum
2. Test extent of /k/ perception for a target phoneme from that continuum
   - Linking Hypothesis: Use "Response Probability" because the behavior is phoneme identification
3. Test the word
   - Linking Hypothesis: Use "Activation" because the behavior is amount of priming as a proxy for prime activation
4. Rinse and repeat (with a new/different input)

Making Ambiguous Phonemes

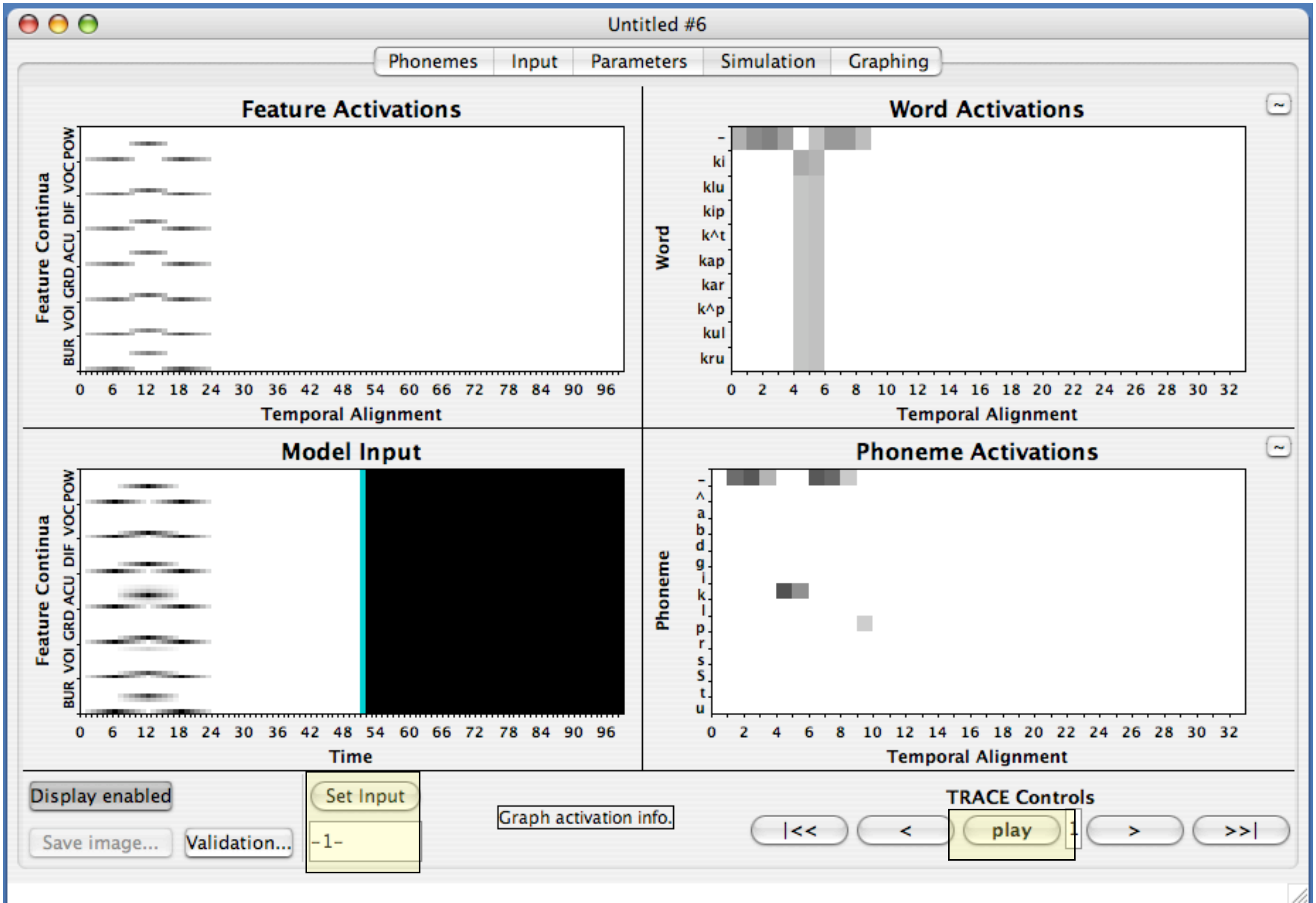
- 6-step /k/-/g/ continuum:



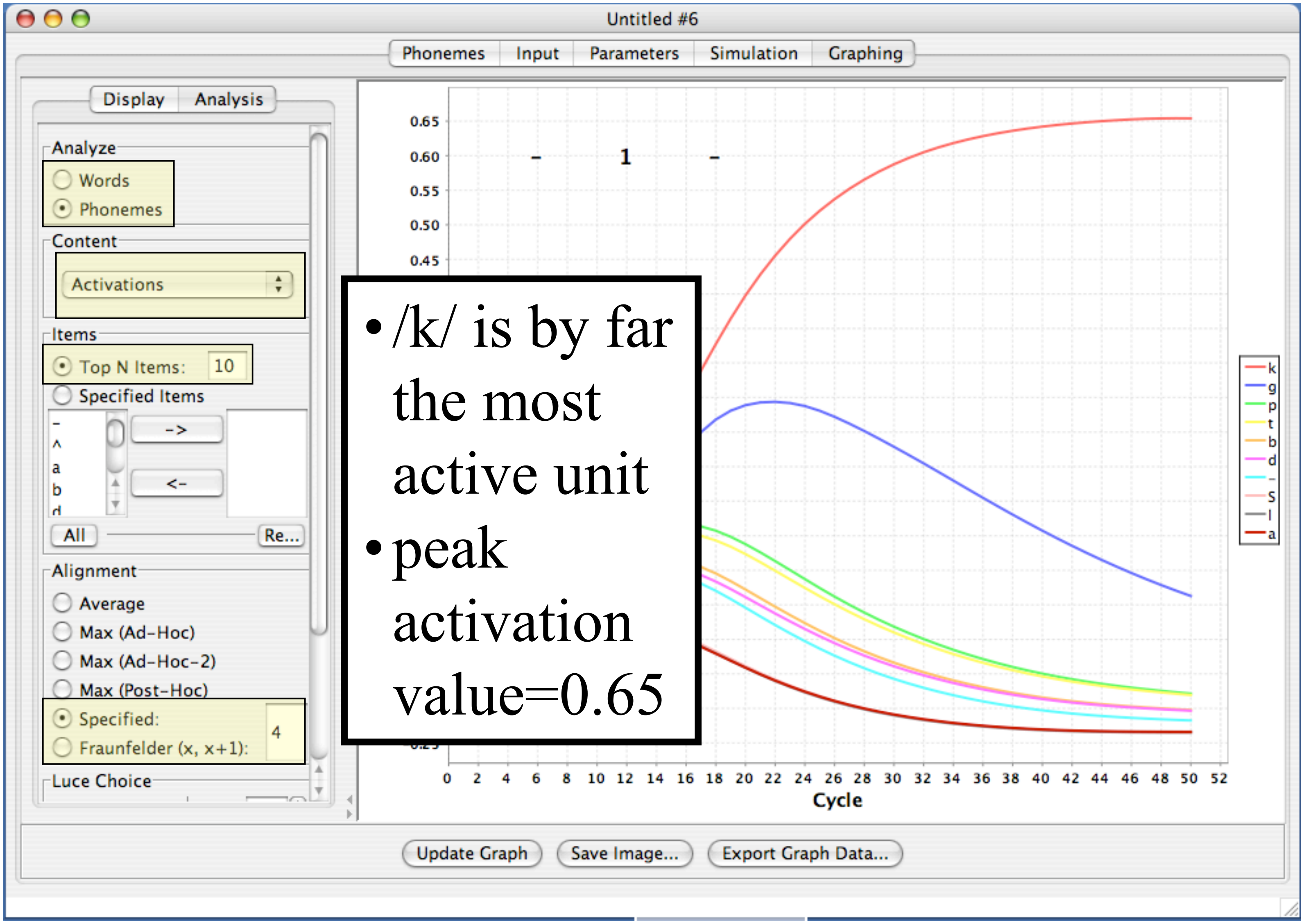
Parameters Tab



Phoneme Identification Pretest: Simulation

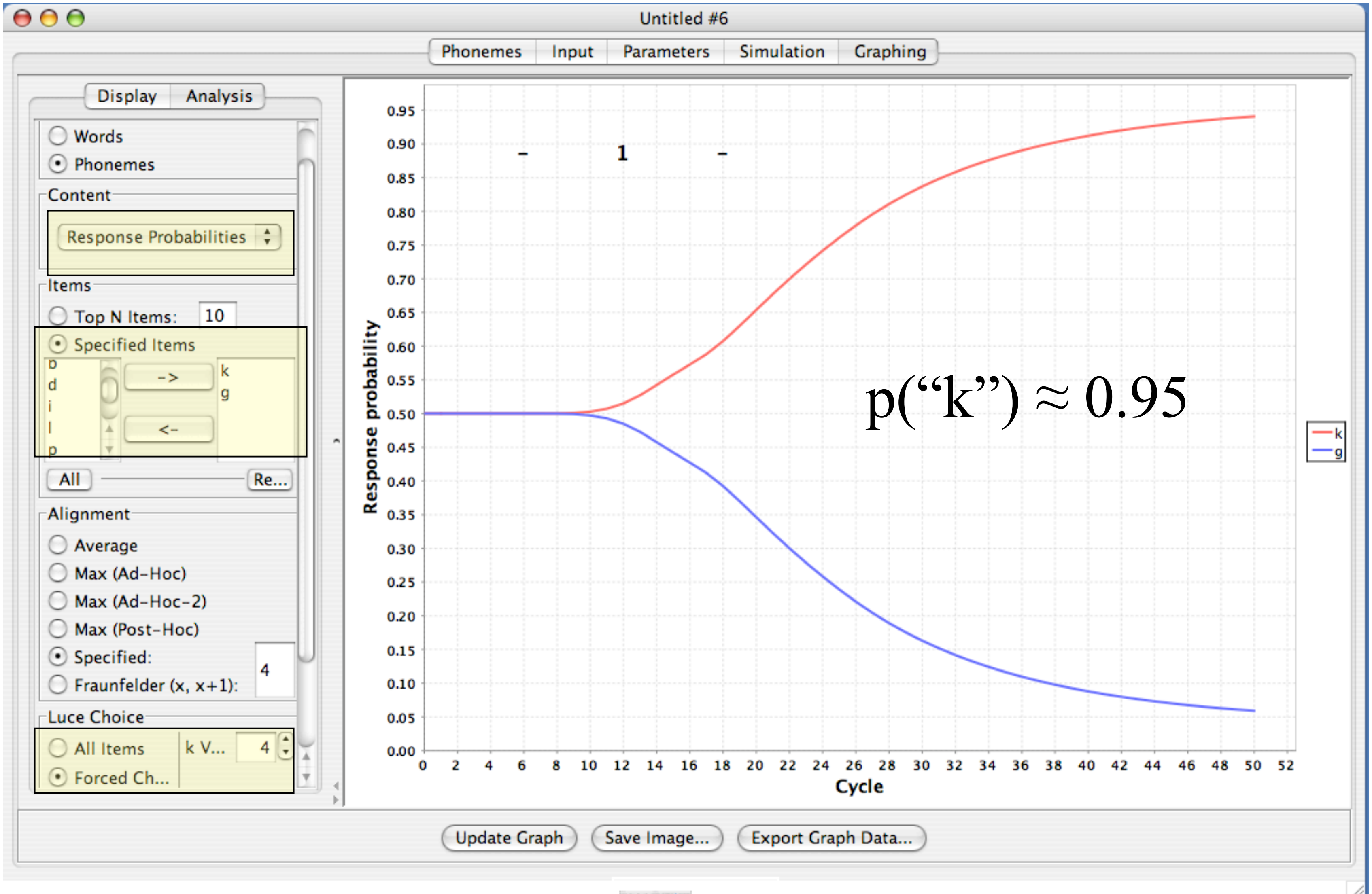


Phoneme Identification Pretest: Activations

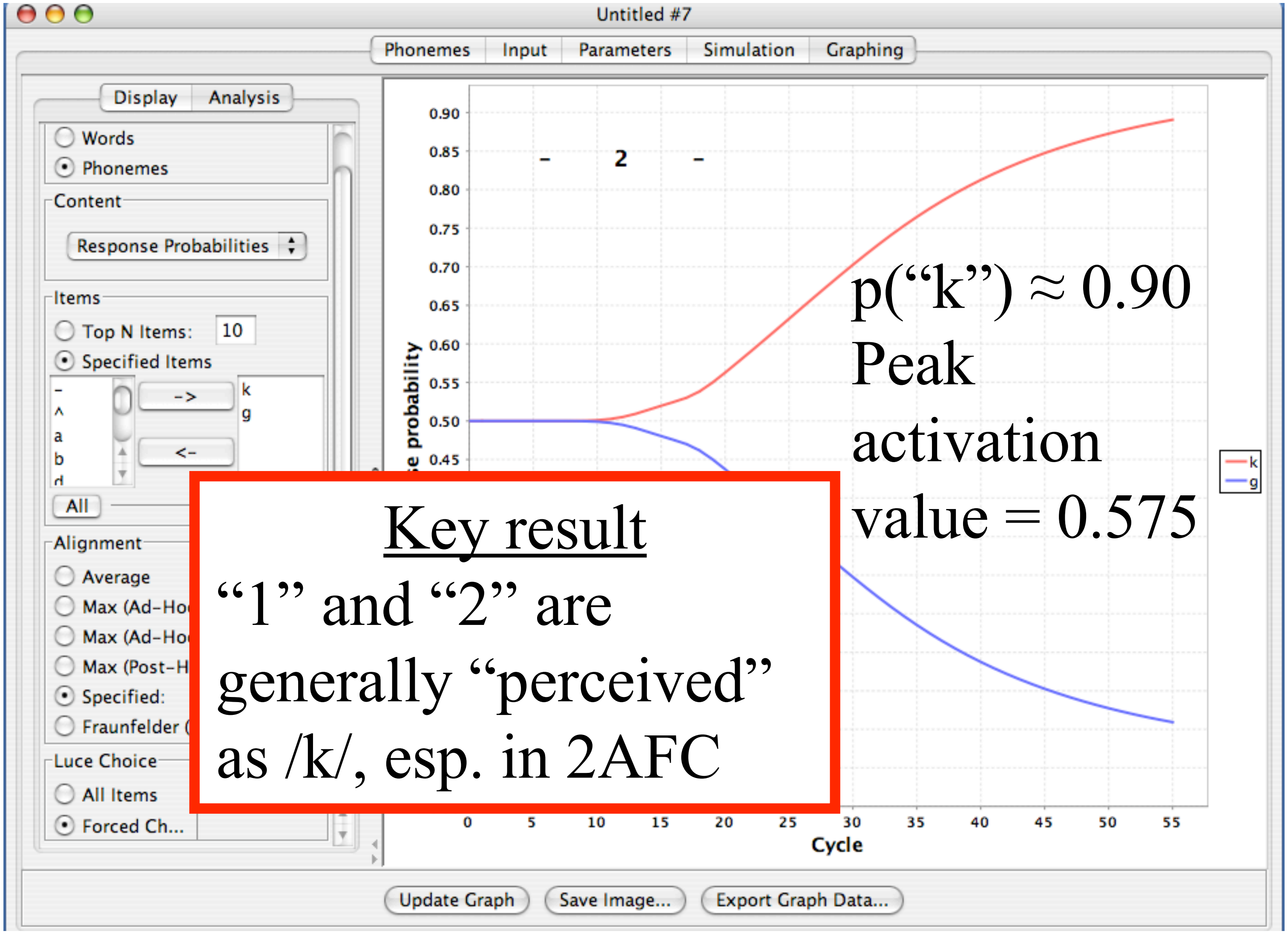


- /k/ is by far the most active unit
- peak activation value=0.65

Phoneme Identification Pretest: 2AFC



Repeat for "2"



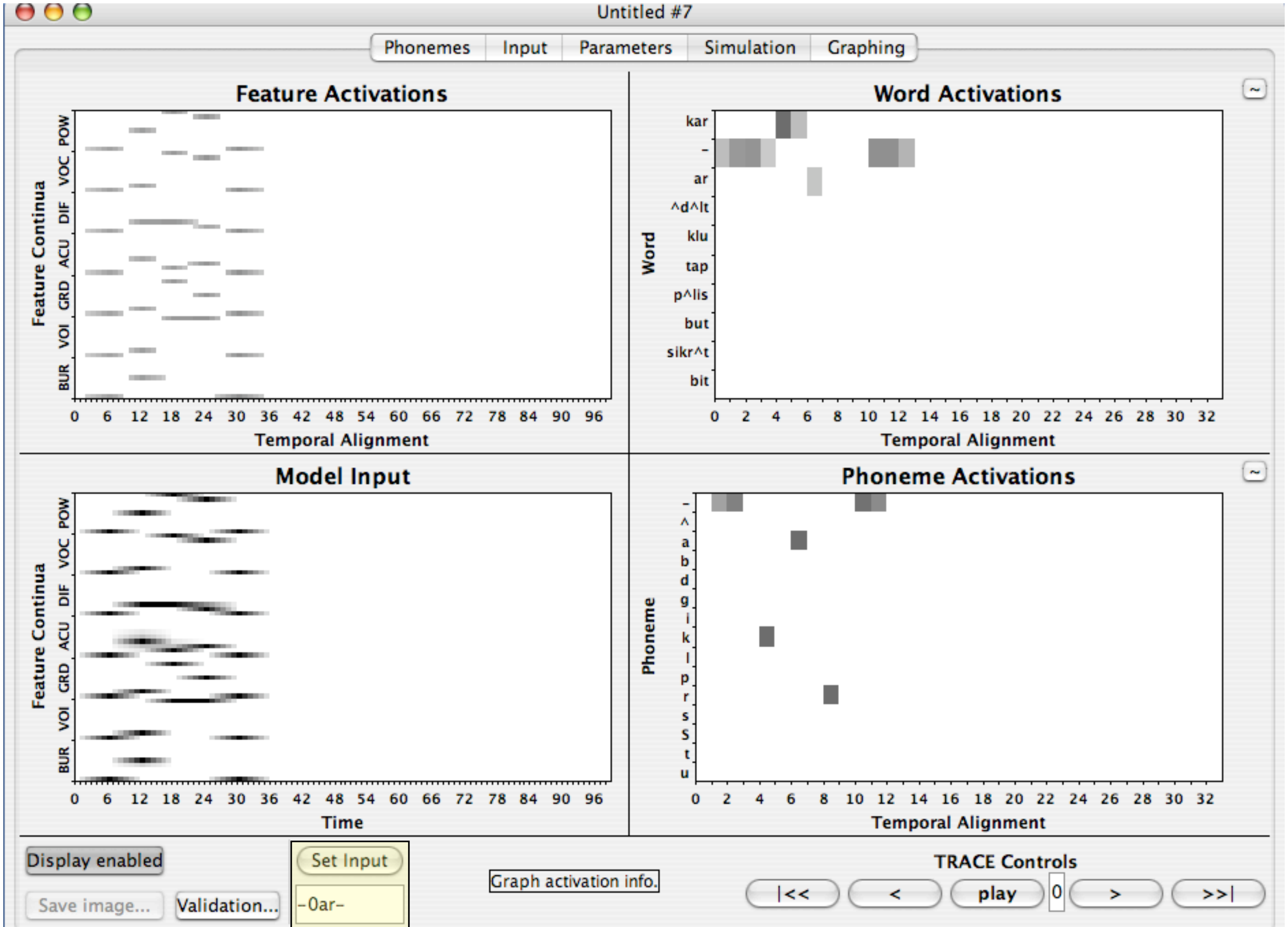
p("k") ≈ 0.90

Peak activation value = 0.575

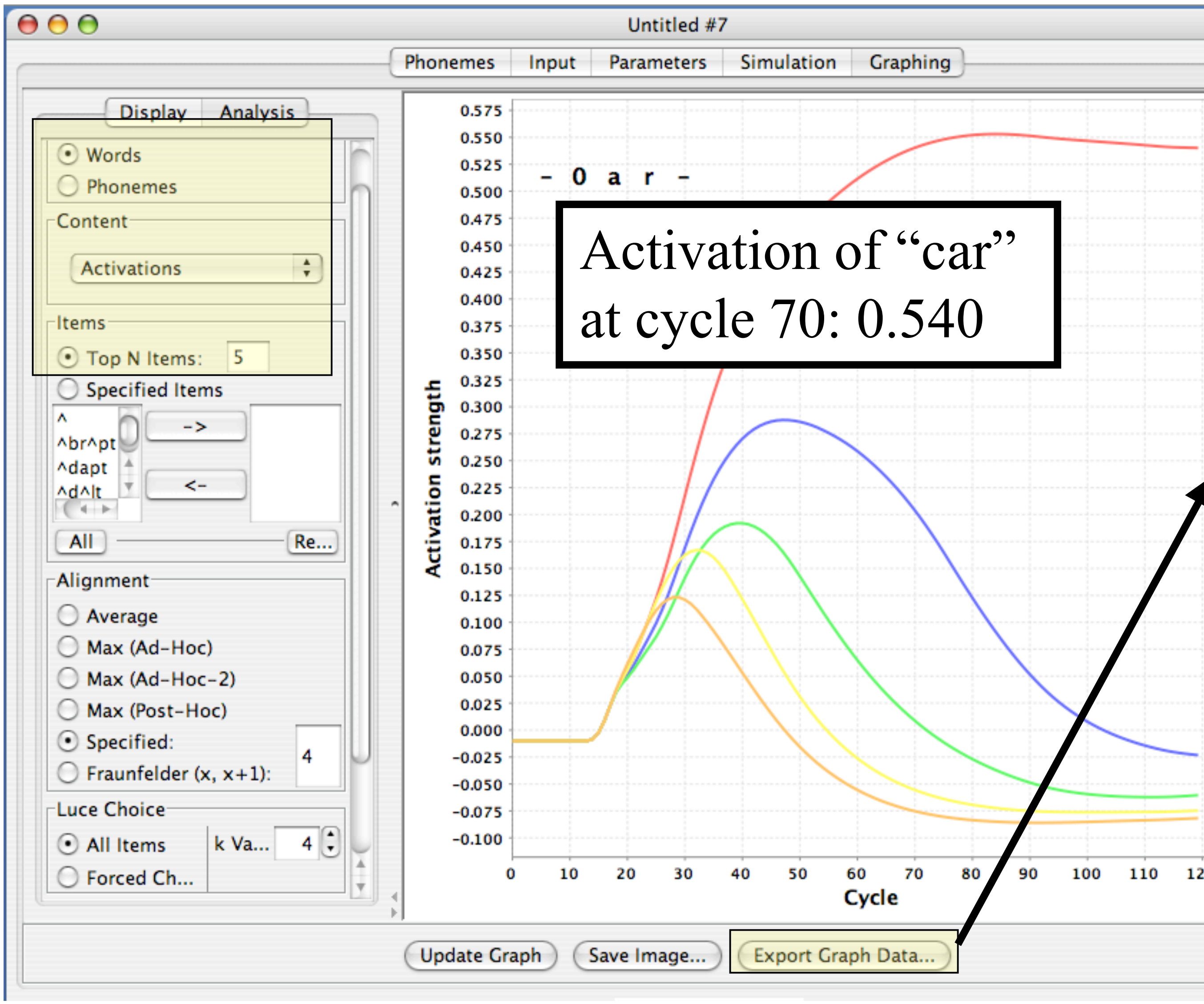
<u>Key result</u>

"1" and "2" are generally "perceived" as /k/, esp. in 2AFC

What is the effect on word activation?



Word Activation



	kar	kard	karp^t	kap	ki
0	-0.01	-0.01	-0.01	-0.01	-0.01
1	-0.01	-0.01	-0.01	-0.01	-0.01
2	-0.01	-0.01	-0.01	-0.01	-0.01
3	-0.01	-0.01	-0.01	-0.01	-0.01
4	-0.01	-0.01	-0.01	-0.01	-0.01
5	-0.01	-0.01	-0.01	-0.01	-0.01
6	-0.01	-0.01	-0.01	-0.01	-0.01
7	-0.01	-0.01	-0.01	-0.01	-0.01
8	-0.01	-0.01	-0.01	-0.01	-0.01
9	-0.01	-0.01	-0.01	-0.01	-0.01
10	-0.01	-0.01	-0.01	-0.01	-0.01
11	-0.010000256758674069			-0.010000256758674069	
12	-0.010003118752305171			-0.010003118752305171	
13	-0.010013028118852775			-0.010013028118852775	
14	-0.008088396273694894			-0.008088396273694894	
15	-0.0020983443521452186			-0.0020983443521452186	
16	0.008614759724045955			0.008614759724045955	
17	0.02246859065363738			0.02240644216247899	
18	0.03587923418111666			0.035208183905449726	
19	0.04603418188451011			0.04377607555174677	
20	0.0556583829010406			0.05111940408819893	
21	0.06672525749476176			0.05937820098561502	
22	0.0796998195283257			0.06906867442077619	
23	0.09323158633564099			0.07884535450625492	
24	0.1073269166606232			0.08883447895320287	
25	0.12214270741248068			0.09927916603413749	
26	0.13869994089191387			0.11125823169989599	
27	0.15661613899174906			0.12432895190978924	
28	0.17616283406860395			0.13879456176103164	
29	0.196650578051807			0.15393732249840666	
30	0.21736955711510164			0.16903065242364138	
31	0.23787026923450189			0.1836537434464538	
32	0.2577718017228779			0.19744914934997645	
33	0.2768570190315639			0.2102492249531822	
34	0.2950571054346173			0.22205597722413584	
35	0.3122012954058548			0.23273180695999307	
36	0.32820818243099525			0.24223057106244708	
37	0.3431462623870572			0.25066863817312335	
38	0.3570824132728202			0.25814135516146947	
39	0.37007581203651935			0.26471913162692645	
40	0.38213998958590245			0.27040320705653764	
41	0.3933028601607386			0.27519924226777165	
42	0.40362031849159513			0.2791389128365157	
43	0.4131493275874683			0.28224911553177434	
44	0.4219916254591987			0.28461117008588116	
45	0.4302228481640724			0.2862824831067285	
46	0.43788565858173367			0.28728894502733354	
47	0.44504524382492777			0.28769464931403943	
48	0.45177265193074134			0.2875773099547773	
49	0.45812825493927745			0.28700627780665006	
50	0.4641559585759167			0.28602956138294283	

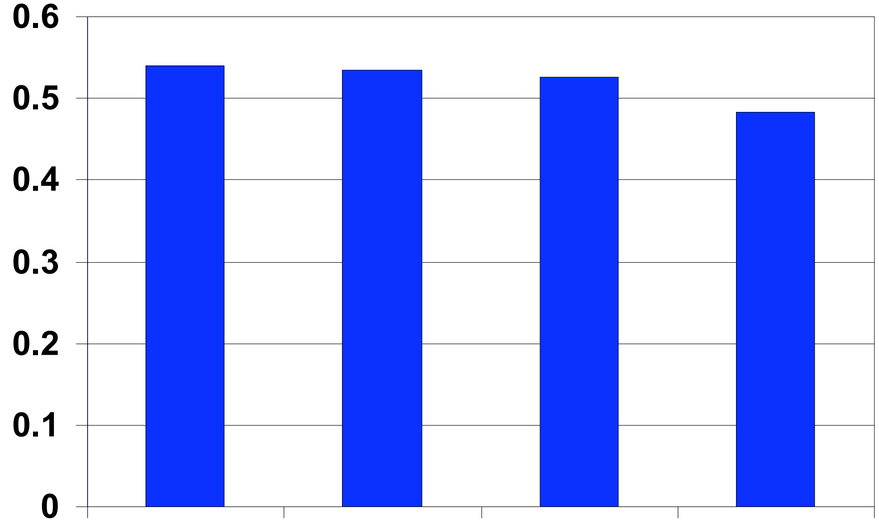
THE NEW SCHOOL

Haskins Laboratories

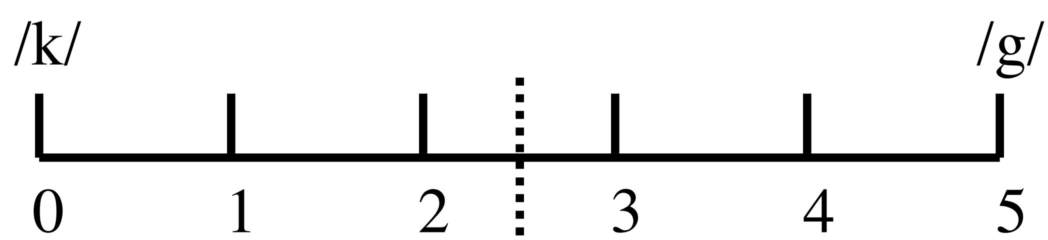
Get "car" activations for: /0ar/, /1ar/, /2ar/, /3ar/

Activation of "car" after 70 cycles of processing:

0.540 0.535 0.525 0.484



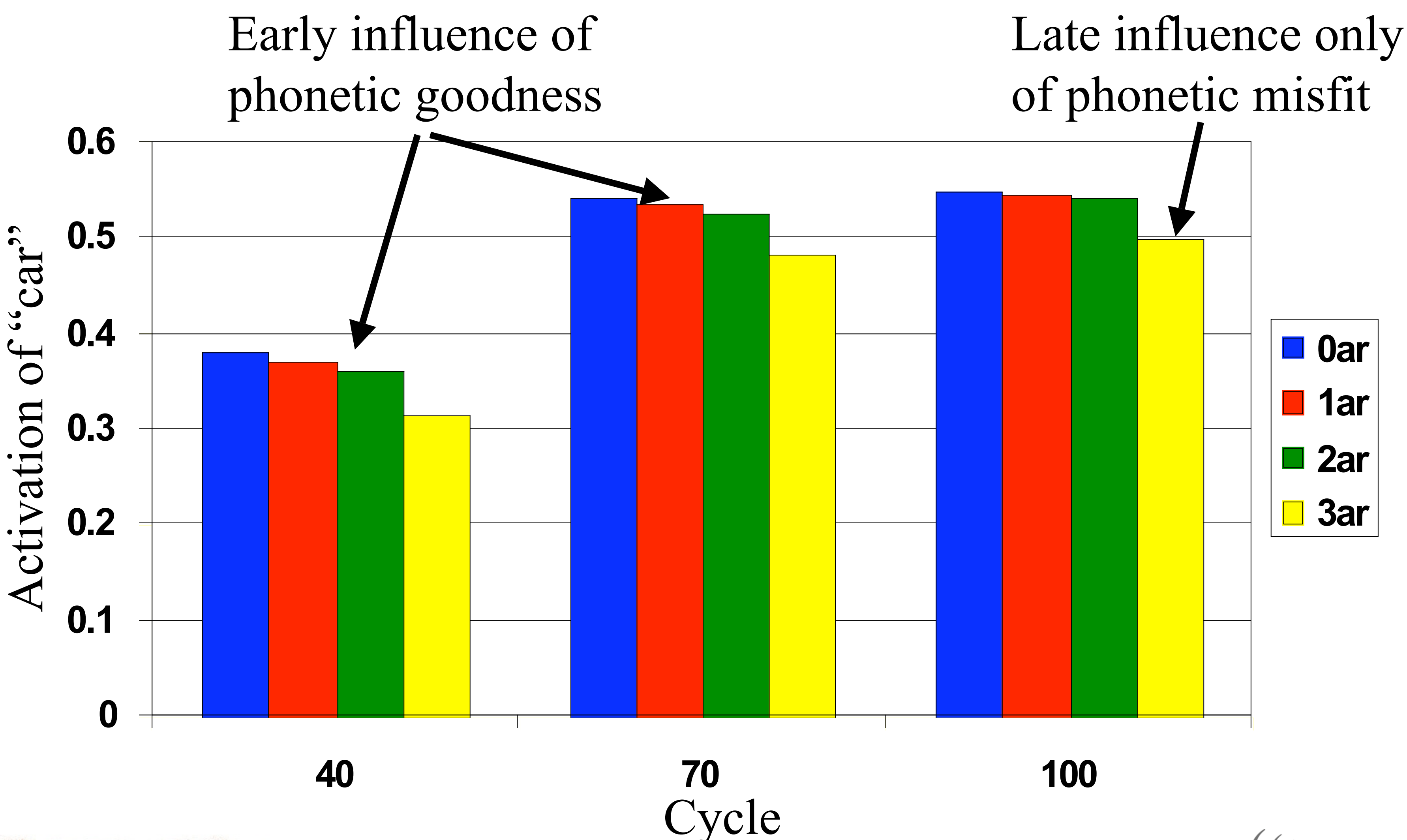
"guard" becomes most active word



THE NEW SCHOOL

Haskins Laboratories

Effect of Processing Time



THE NEW SCHOOL

Haskins Laboratories

Why does the model exhibit this pattern?

- Cascading processing
  - Activated phoneme units begin activating lexical units before phoneme processing is "done"
- Lateral inhibition
  - Initially /k*/ activates /k/ and partially activates /g/
  - /k/ and /g/ compete through lateral inhibition
  - Eventually /k/ wins competition

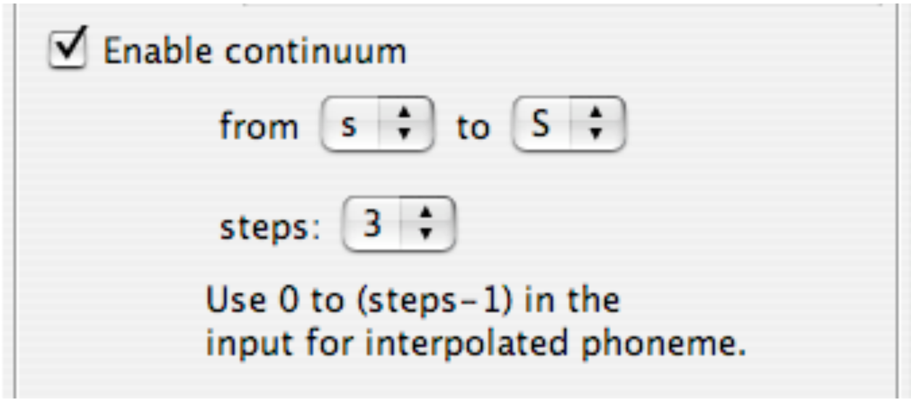
3. Lexical Effect on Phoneme ID

- Basic behavioral finding(s) (Ganong, 1980)
  - The same ambiguous fricative between /s/ and /S/ is heard as /s/ at the end of /b^_/ and as /S/ at the end of /r^_/
- Interpretation and linking to model
  - Lexical information directly influences phoneme processing
  - Excitatory feedback connections from word units to phoneme units
  - Ambiguous phoneme activation will be biased to be lexically consistent

Simulation steps

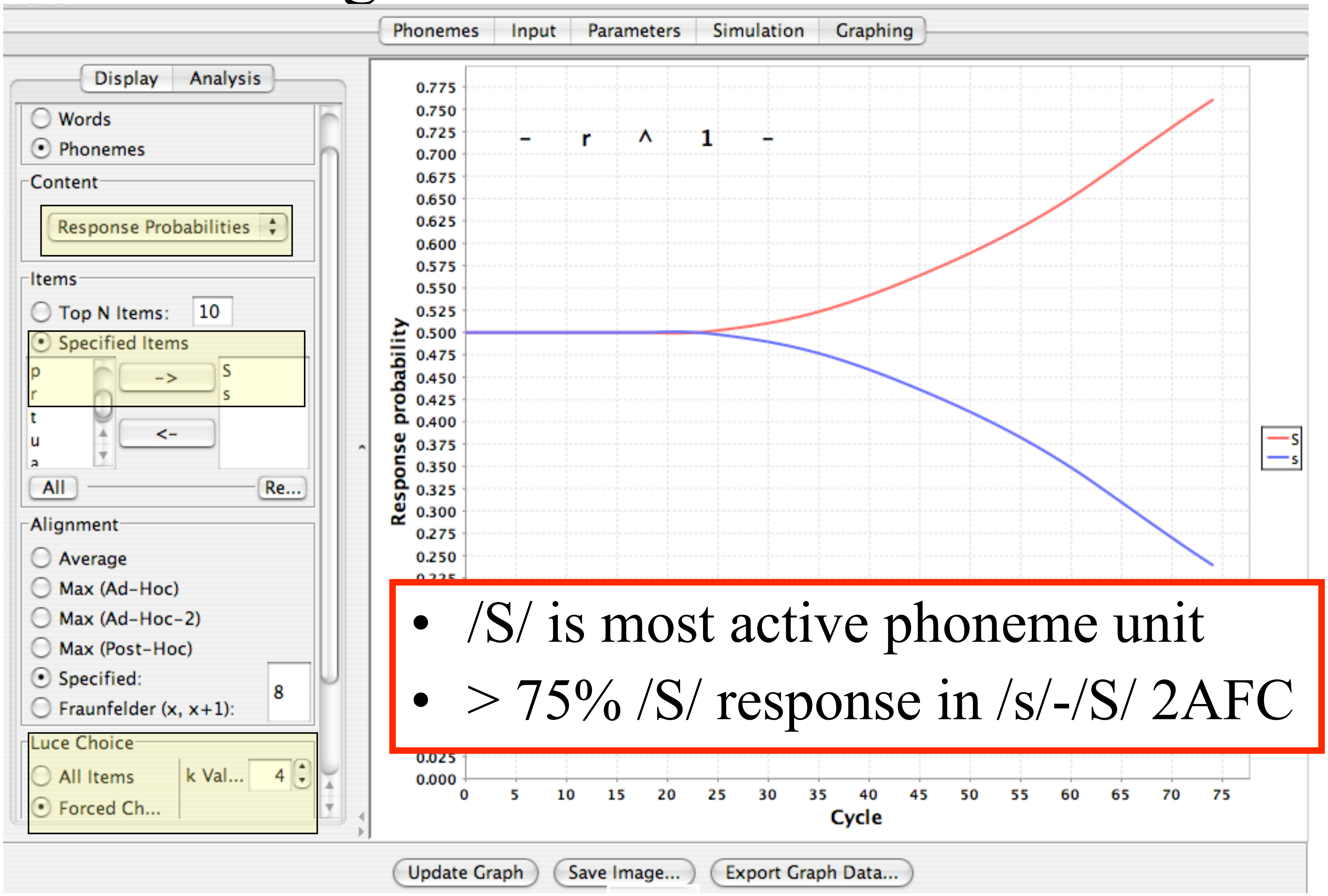
1. Create ambiguous phoneme

Parameters tab



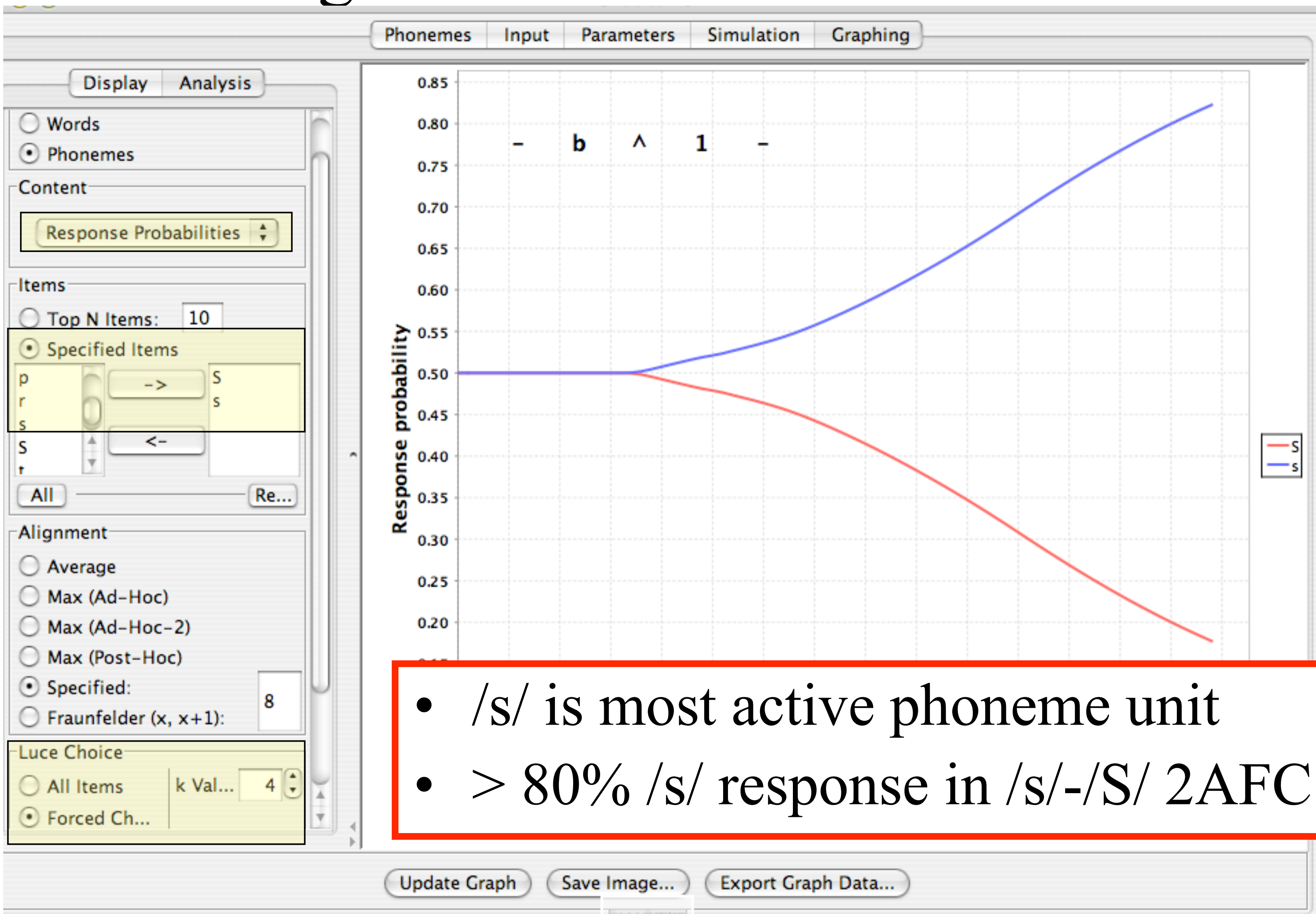
2. Test ambiguous phoneme in opposite lexical contexts
   i. /S/-bias: /r^_/
   ii. /s/-bias: /b^_/
3. Control test: ambiguous phoneme in isolation

How is Ambiguous Phoneme Perceived in /r^1/

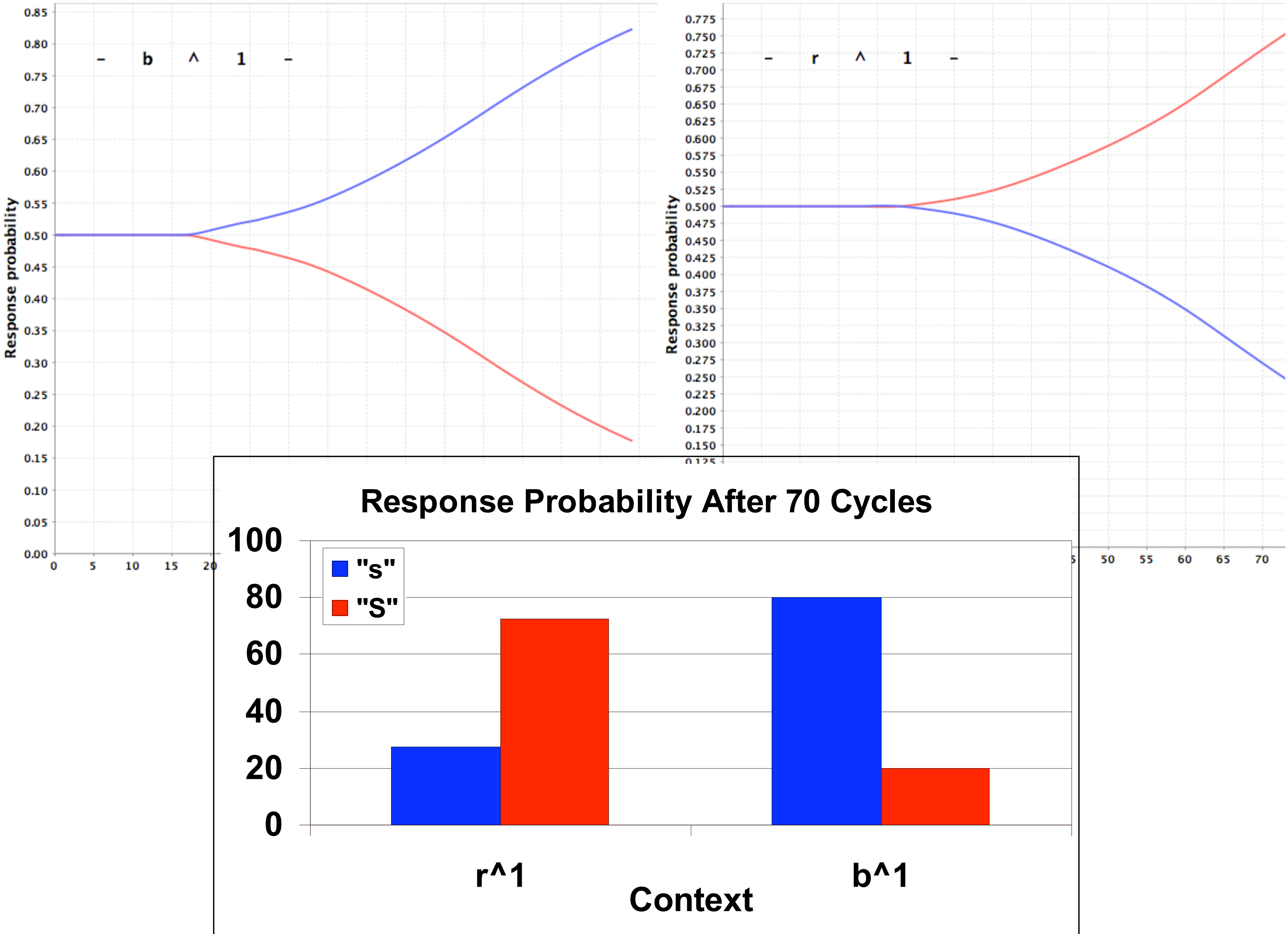


- /S/ is most active phoneme unit
- > 75% /S/ response in /s/-/S/ 2AFC

How is Ambiguous Phoneme Perceived in /b^1/



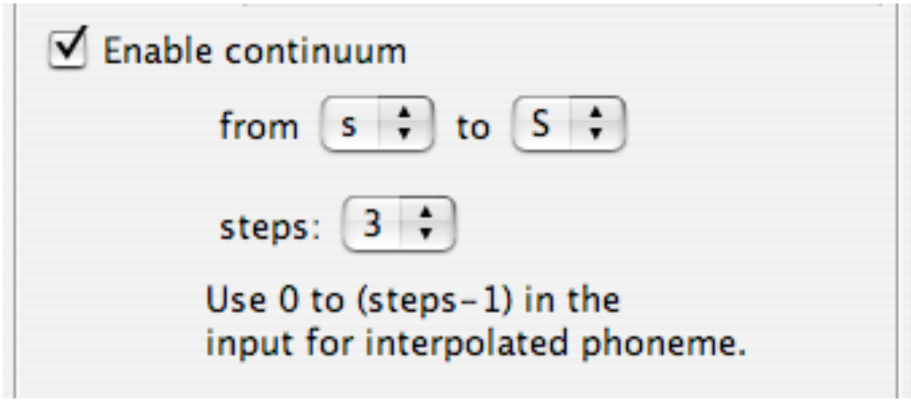
- /s/ is most active phoneme unit
- > 80% /s/ response in /s/-/S/ 2AFC


– b ^ 1 –
Response probability
– r ^ 1 –
Response probability
Response Probability After 70 Cycles
100
80
60
40
20
0
"s"
"S"
r^1
b^1
Context

Simulation steps

✓ Create ambiguous phoneme

Parameters tab



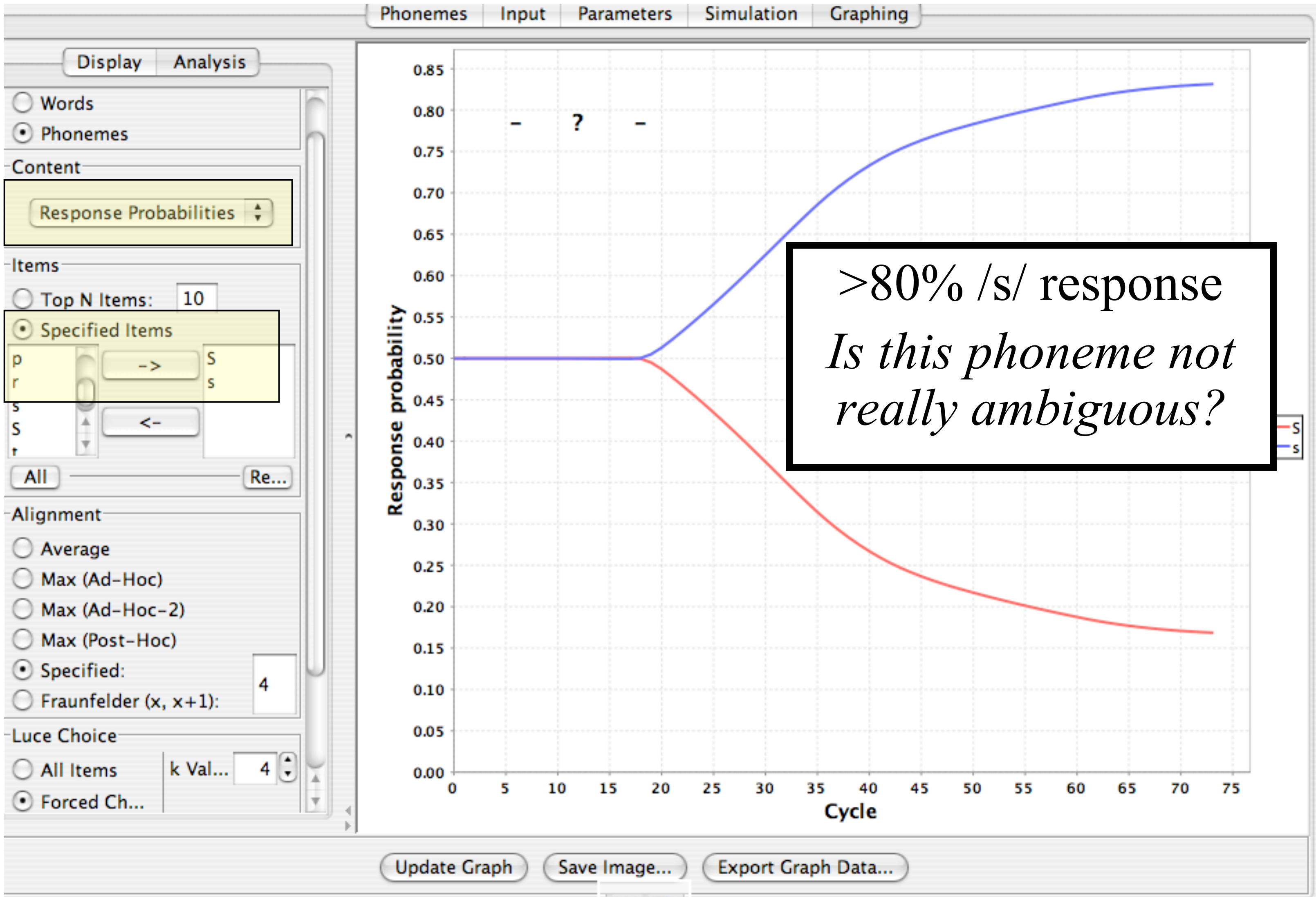
✓ Test ambiguous phoneme in opposite lexical contexts

i. /S/-bias: /r^_/

ii. /s/-bias: /b^_/

3. Control test: ambiguous phoneme in isolation

Control Test: Ambiguous Phoneme in Isolation



Turning off the lexicon: Two ways

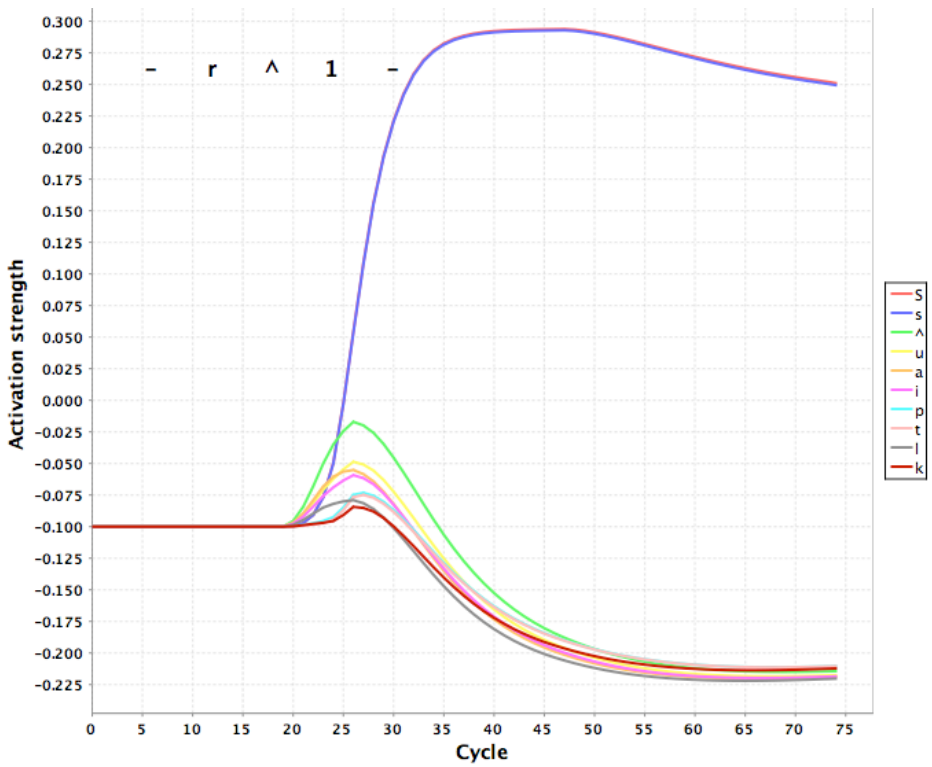
Parameters tab…

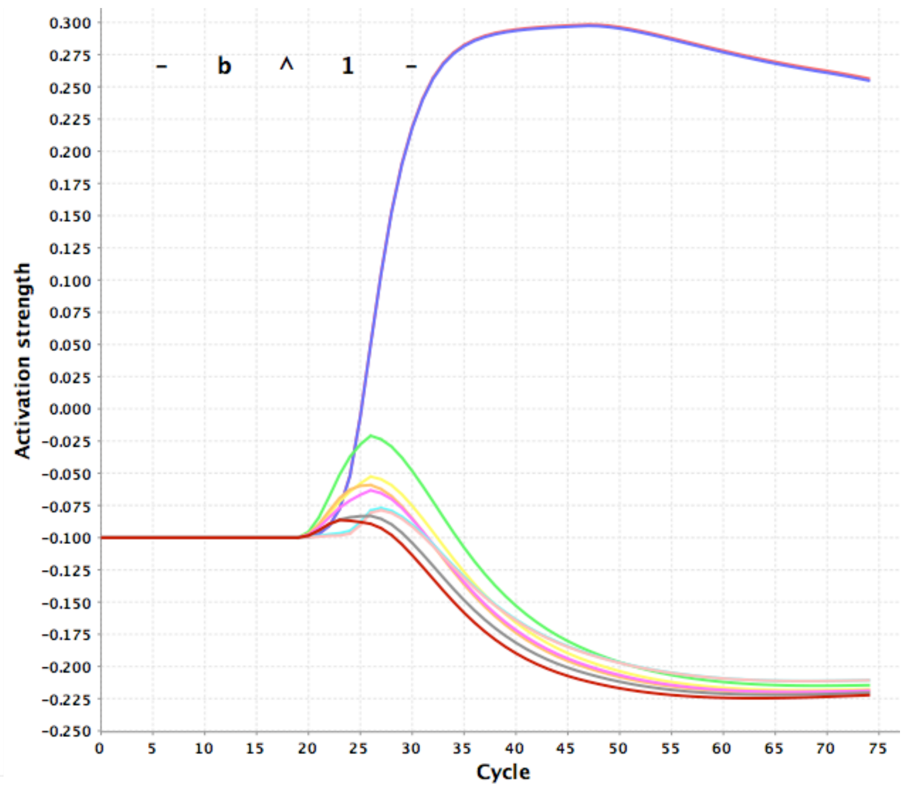
Parameters

Parameter	Value	Function	Default	Notes
Comment		-		
User		-		
Date		-		
aLPHA[if]	1	-	1	Input-Feature weights
aLPHA[fp]	0.02	-	0.02	Feature-Phoneme weights
aLPHA[pw]	0.05	-	0.05	Phoneme-Word weights
aLPHA[pf]	0	-	0	Phoneme-Feature weight
aLPHA[wp]	0.03	-	0.03	Word-Phoneme weights
GAMMA[f]	0.04	-	0.04	Feature-layer inhibition
GAMMA[p]	0.04	-	0.04	Phoneme-layer inhibition
GAMMA[w]	0.03	-	0.03	Word-layer inhibition
DECAY[f]			0.01	Feature decay
DECAY[p]			0.03	Phoneme decay
DECAY[w]			0.05	Word decay
REST.F			-0.1	Feature resting activation
REST.P	-0.1	-	-0.1	Phoneme resting activation
REST.W	-0.01	-	-0.01	Word resting activation
Input Noise (SD)	0	-	0	
Stochasticity (SD)	0		0	McClelland: 0.02
Attention	1	-	1	Lexical gain
Bias	0	-	0	Lexical bias
Learning Rate	0	-	0	Coming soon... (ft->ph lear...
spreadScale	1	-	1	Scales FETSPREADs
min	-0.3	-	-0.3	Minimum activation
max	1	-	1	Maximum activation

Set one of these values to 0.0

Does turning off the lexicon eliminate the lexical effect?

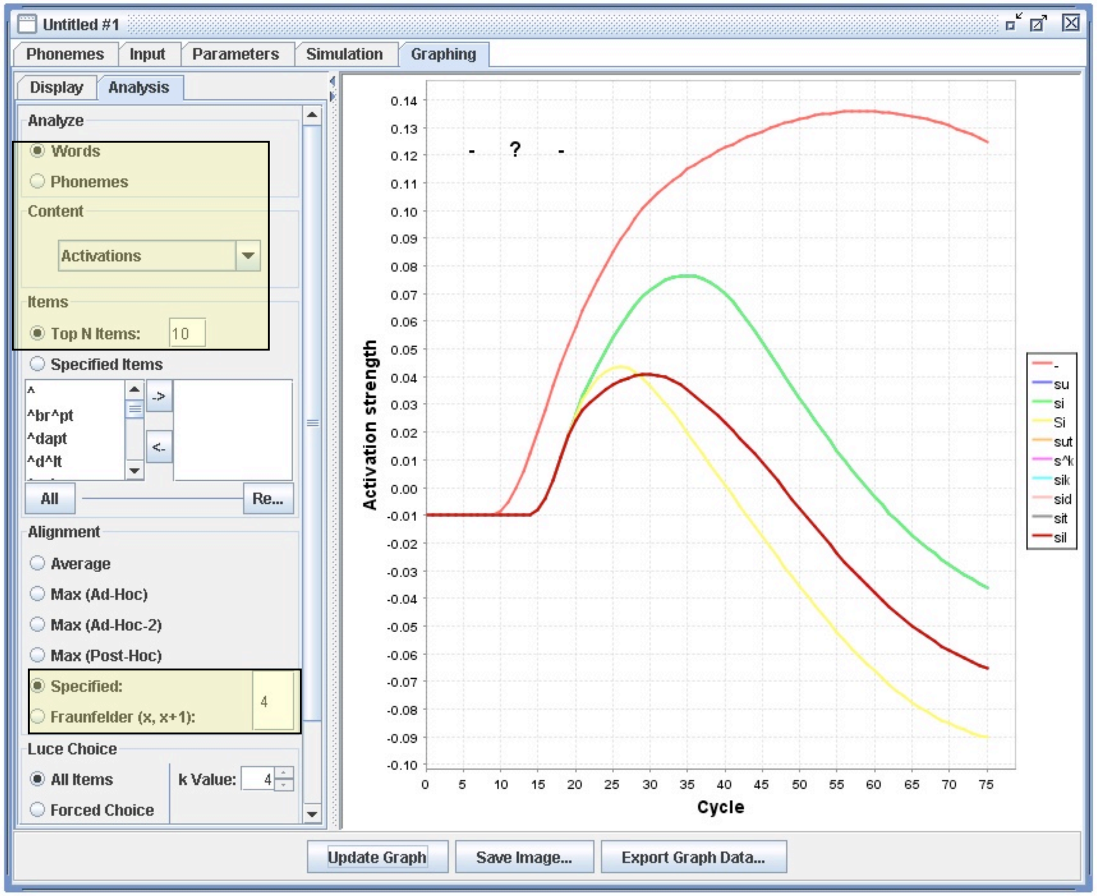




Also eliminates the /s/ advantage?

What does that say about its cause?

Word activation to -?- input



(Turn lexicon back on)

10 most active words:

\-
sue
see
she
suit
suck
seek
seed
seat
seal

- Partial activations matter: Neighborhoods (gangs) of words can influence phoneme perception
- Context also affects preceding phonemes

Test your jTRACE skills: Try one of these exercises…

1. Is the Ganong effect the same for word-initial and word-final phonemes?
2. Turn on frequency and test high vs. low frequency words
   - Is there a difference in speed of recognition?
   - Is there a frequency analog to the lexical effect on phoneme perception?
3. What are the differences in processing of short and long words?
   - Which are recognized faster?
   - What happens when they compete?
   - Do they differ in terms of effects on phoneme perception?