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

THE NEW SCHOOL

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 2

- Motivation for jTRACE
- Tour of jTRACE

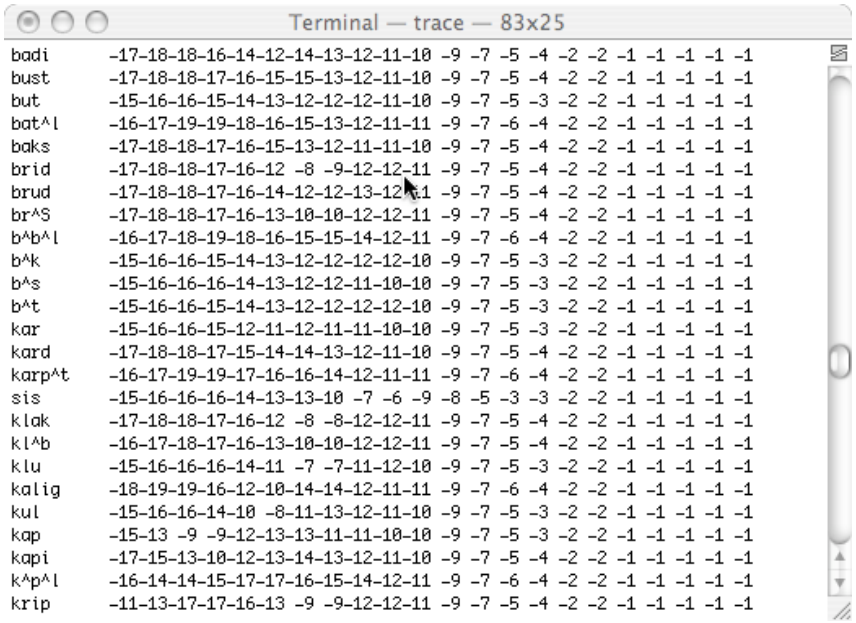
THE NEW SCHOOL

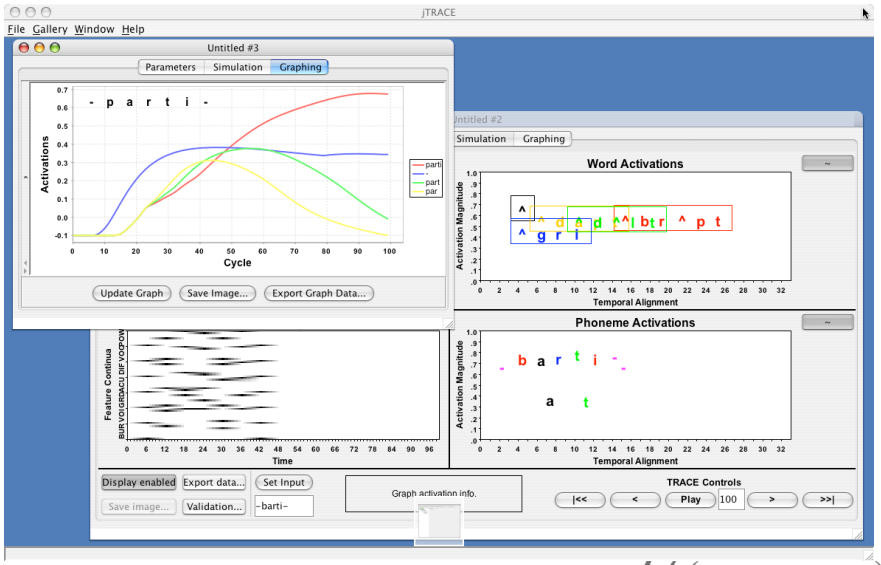
Haskins Laboratories

Why jTRACE?

- Original TRACE program was written in C. Requires computer savvy, time.
- We created jTRACE (with Java) to make TRACE simulation accessible to average PC users.
- Simulation, scripting and analysis can all be done within jTRACE. Data can be dumped for further processing in SPSS, MS Excel, Matlab, etc.
- The main advantage of jTRACE is its user-friendly interface.

cTRACE → jTRACE





THE NEW SCHOOL

Haskins Laboratories

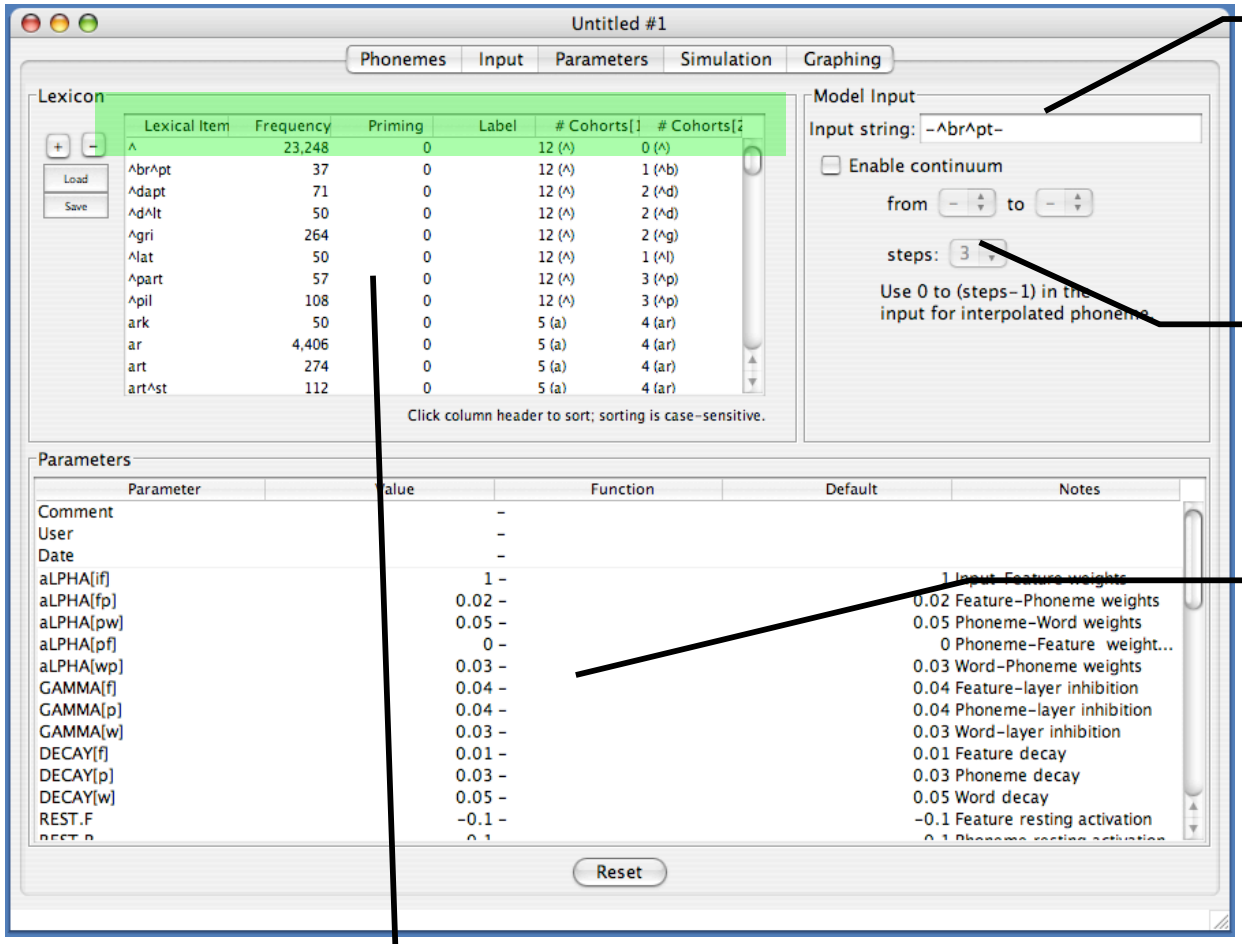
cTRACE = jTRACE ??

- We validated jTRACE against TRACE by running the same simulations in both models and comparing the outputs.
- In the *worst* case, the two versions of TRACE differed by 0.007%.

jTRACE basics

- A simulation is treated like a document.
- Each simulation has a parameter set, a lexicon, an input string, a simulation display panel, a graph panel, and analysis settings.
- **Parameters panel** - set parameters, lexicon, input, ambiguous phonemes, etc.
- **Simulation panel** - run the simulation and watch progress as an animation.
- **Graphing panel** - graph activations and response probabilities of individual word/phoneme units.
- **Input panel** - setup and visualize the model's input
- **Phoneme panel** - edit, create phoneme set

Parameters panel



input string

ambiguous phonemes

parameters table

lexicon table

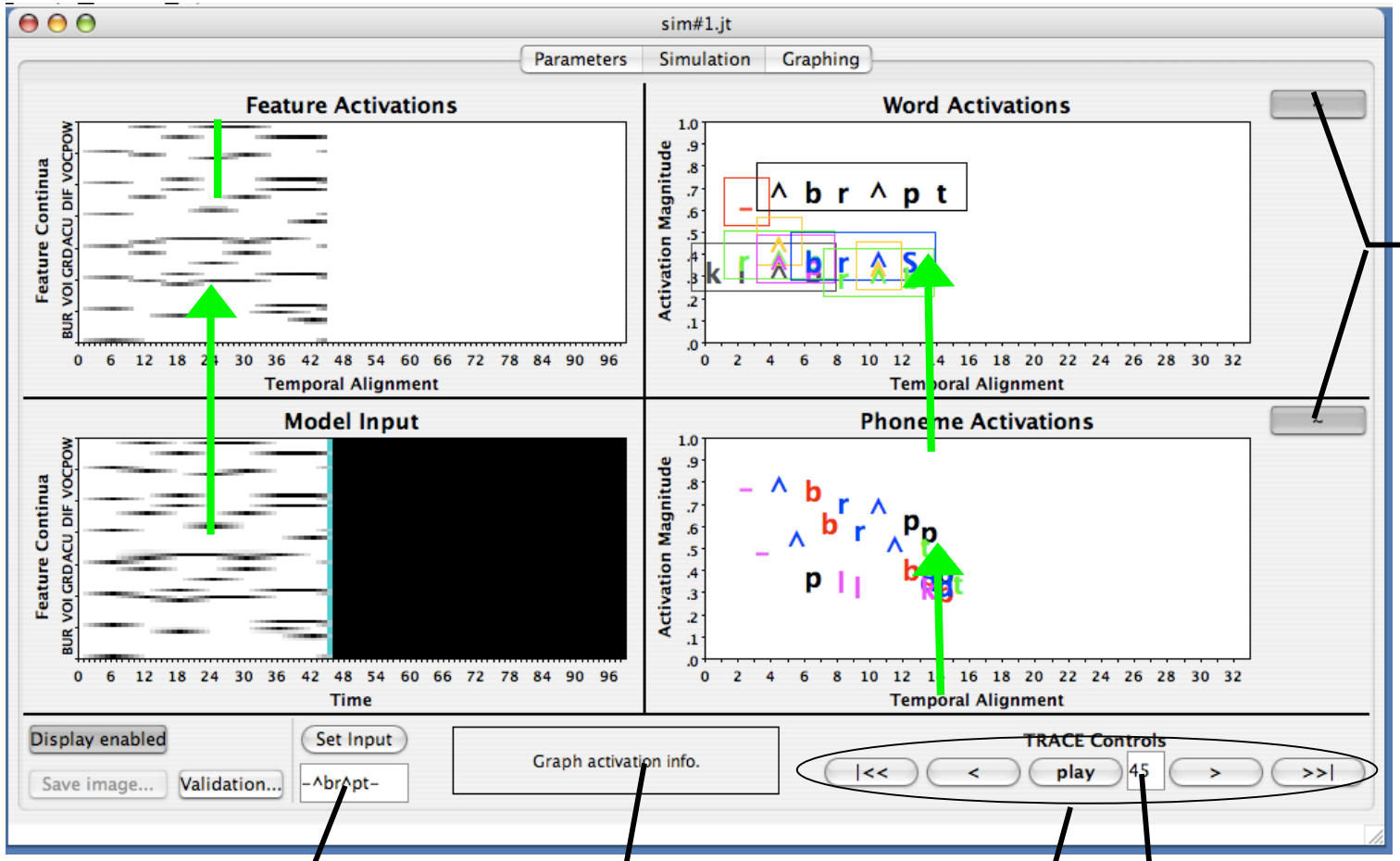
Lexicon notes

- Composition of the lexicon is key to word competition and attendant processes.
  - Trade-off between complex and simple lexicons.
- Default lexicon has 213 words. A larger lexicon of 901 words is included, plus some others.
- Lexicons can be edited in the table or written by hand in a file using a simple XML markup.

```
<lexeme><phonology>^br^pt</phonology><frequency>37</frequency></lexeme>
```

- Phoneme roster includes 14+1 phonemes: `a ^ i u p b t d k g s S r l`; and `-` is the *silence* phoneme.
  - It is possible to add/edit phonemes.

Simulation panel



toggle view

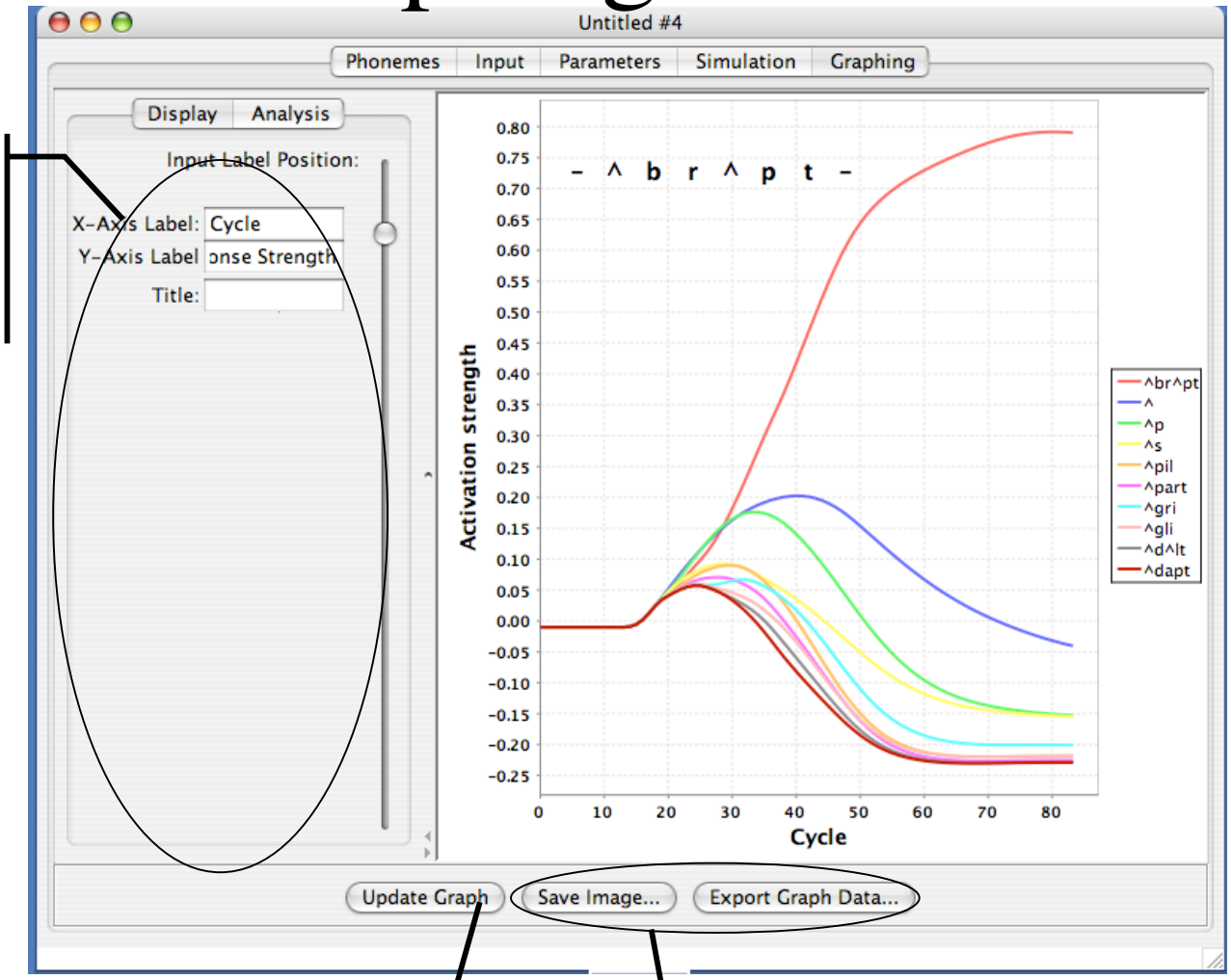
change input

query cell

simulation controls

cycle #

Graphing Panel



analysis & display settings

update button

export data

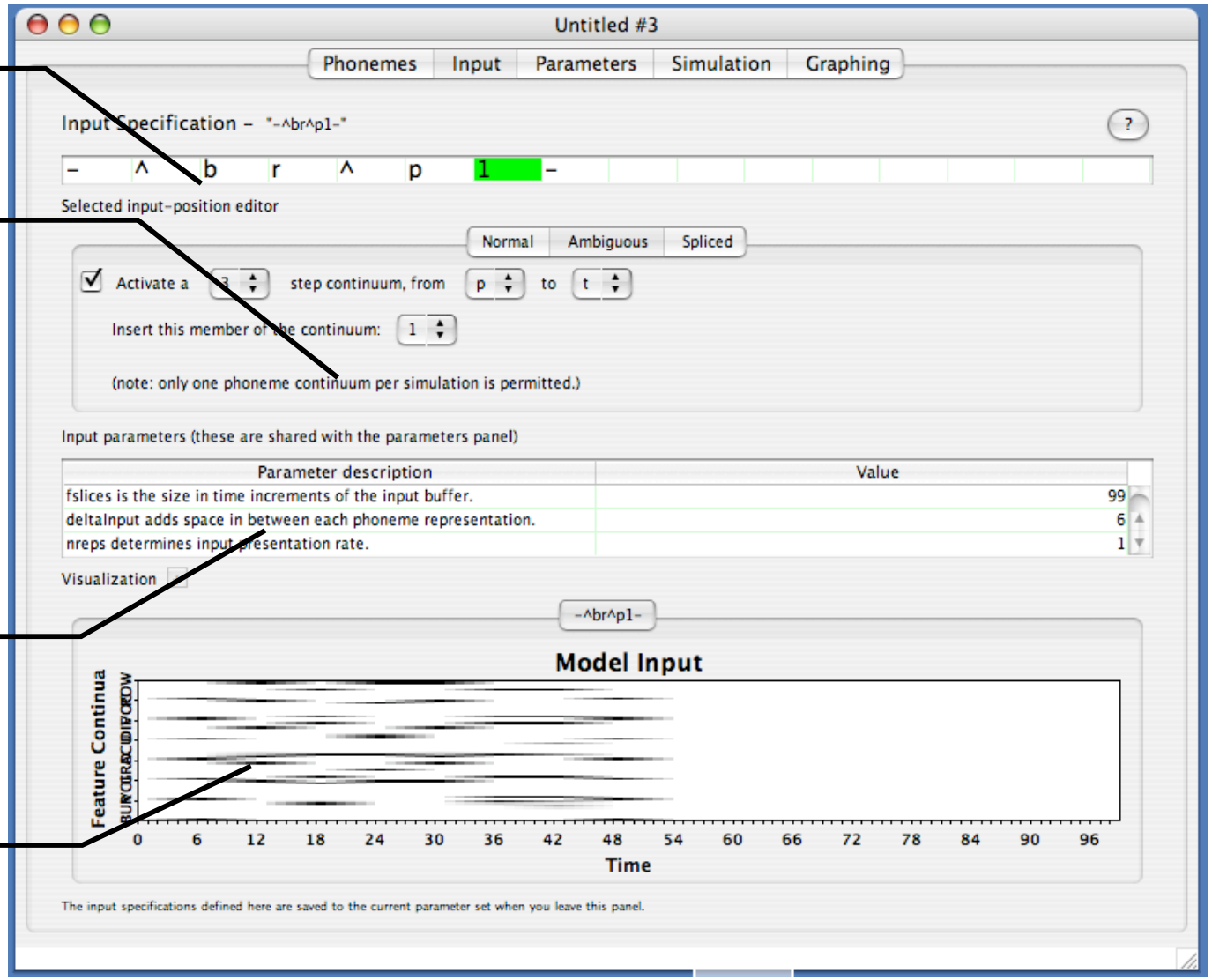
Input Panel = training wheels

letter-by-letter spec

setup the highlighted phoneme using simple menus (normal, ambiguous, spliced)

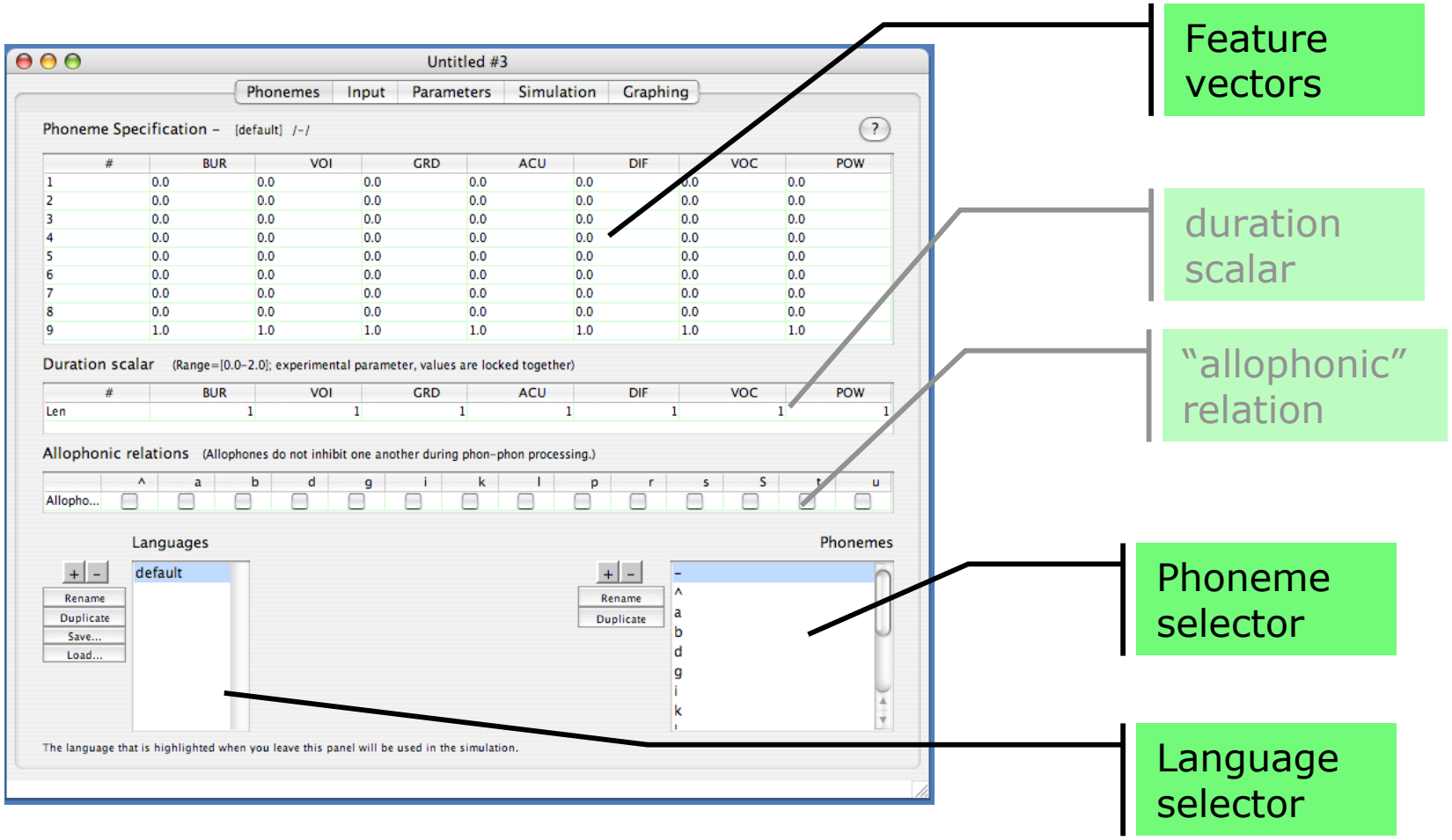
model parameters re: input

input visualizer



Phoneme panel (experimental)

- Edit phoneme specs; custom phonologies.



First jTRACE simulation

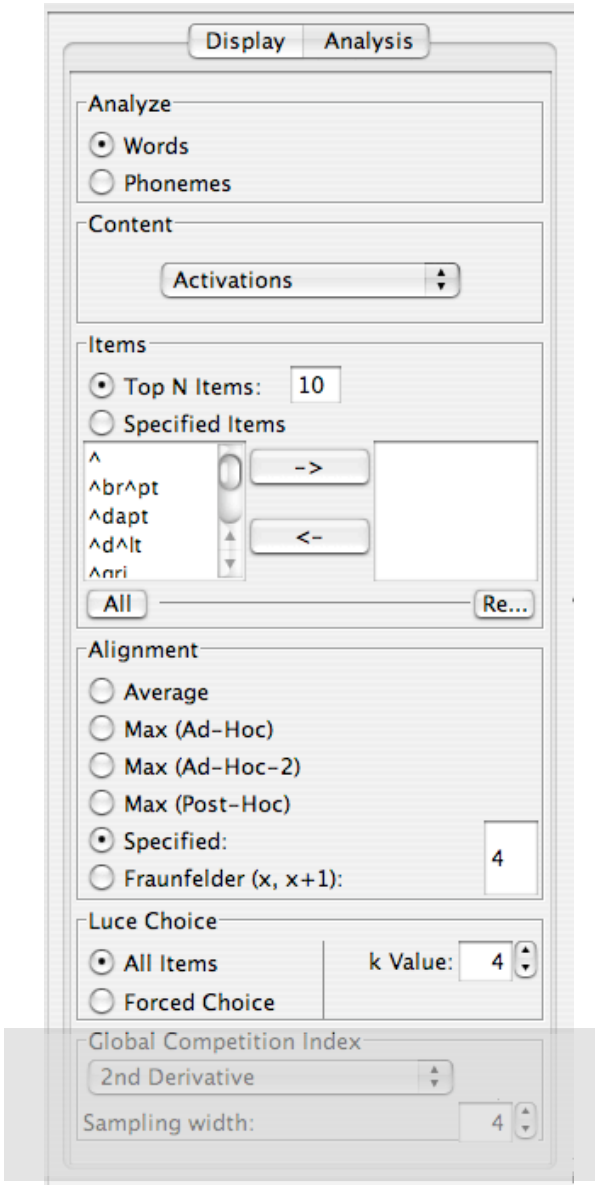
- File > new model
- Parameters tab > choose a word from the lexicon > type it into the input string field
- Simulation tab > play > the "~" button changes visualization > stop at cycle 70
- Graphing tab > save image... > file-name.png
- File menu > save as... > file-name.jt

How to *read* a simulation

- Open a new simulation, enter "-tik^p-" as input string
- Simulation panel, run for 80 cycles
- Graphing panel: Top N Items=5, Alignment=Max (post-hoc)
- Press update graph button
- What do we see?
  - Recognition of words in sequence
  - Lexical competition amongst cohort of k^p
  - Monotonic activation, peak, plateau, decay
  - Other observations?

Graph analysis options

- Analyze
  - words / phonemes
- Content
  - activations / response probailities / competition index
- Items
  - top N / specific items
- Alignment
  - avg / **max...** / **specified** / Frauenfelder
- Luce Choice Rule (LCR)
  - choice: normal / forced
  - k value

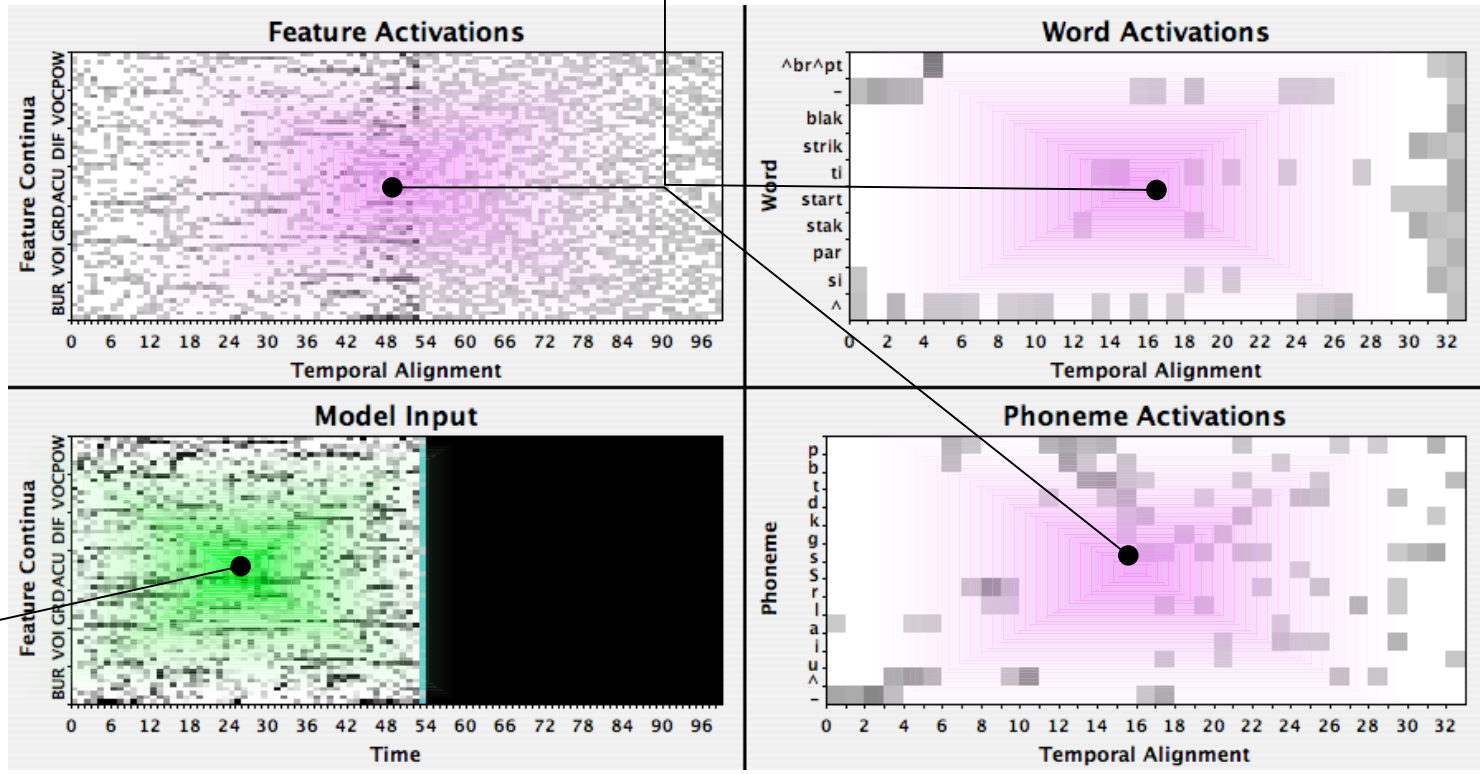


Some useful parameters

- Main TRACE parameters
  - Alpha (excitation, feedforward & feedback)
  - Gamma (within-layer inhibition)
  - Rest (resting level activation)
  - Decay (rate of unit decay)
- Extensions of the model
  - Input noise, stochasticity
  - Frequency effects
  - Priming effects
  - Attention - gain, bias

Using noise & stochasticity

Stochasticity
(applied every cycle)



Input noise (applied once)

Using frequency

1. The TRACE lexicon must contain frequency values.
2. Activate desired frequency implementation in the parameters table.

Lexical Items	Frequency
^	23,248
^br^pt	37
^dapt	71
^d^lt	50
^gri	264
^lat	50
^part	57
^pil	108
ark	50
ar	4,406
art	274
art^st	112

Parameters

Parameter	Value	Default	Notes
spreadScale	1	1	Scales FETSPREADs
min	-0.3	-0.3	Minimum activation
max	1	1	Maximum activation
frq resting levels	0	0	Dahan et al.: 0.06
frq p->w wts	0.13	0	Dahan et al.: 0.13
frq post-act	0	0	Dahan et al.: c=15
FETSPREAD.pow	6	6	Power feature spread
FETSPREAD.voc	6	6	Vocalic feature spre...
FETSPREAD.dif	6	6	Diffuse feature spre...

Priming

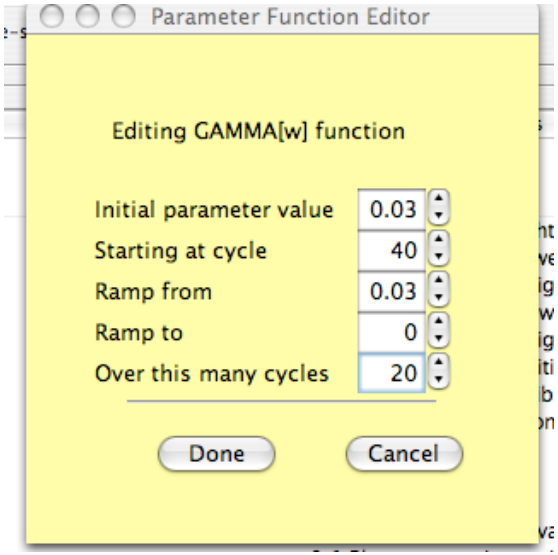
- Exactly the same as using frequency; see previous slide.

Attention

- Lexical bias - negative input to all words
- Lexical gain - responsiveness of lexical units to their input.

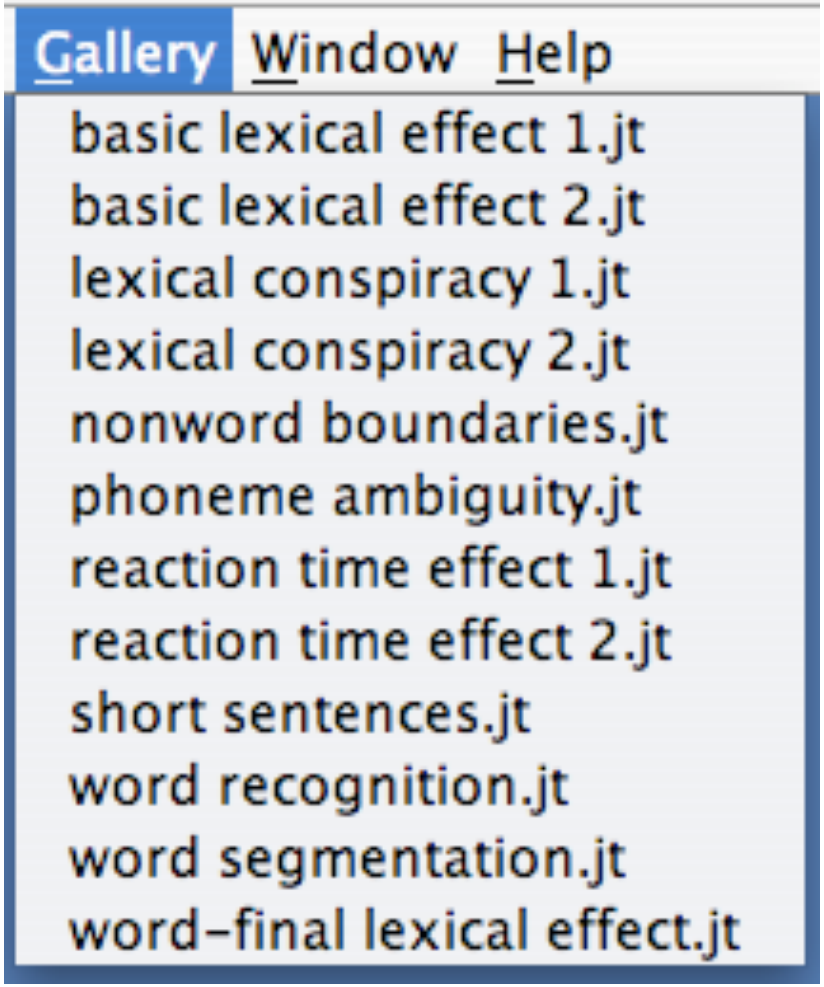
Parameter functions

- Use a linear function to ramp a function on or off during a simulation.



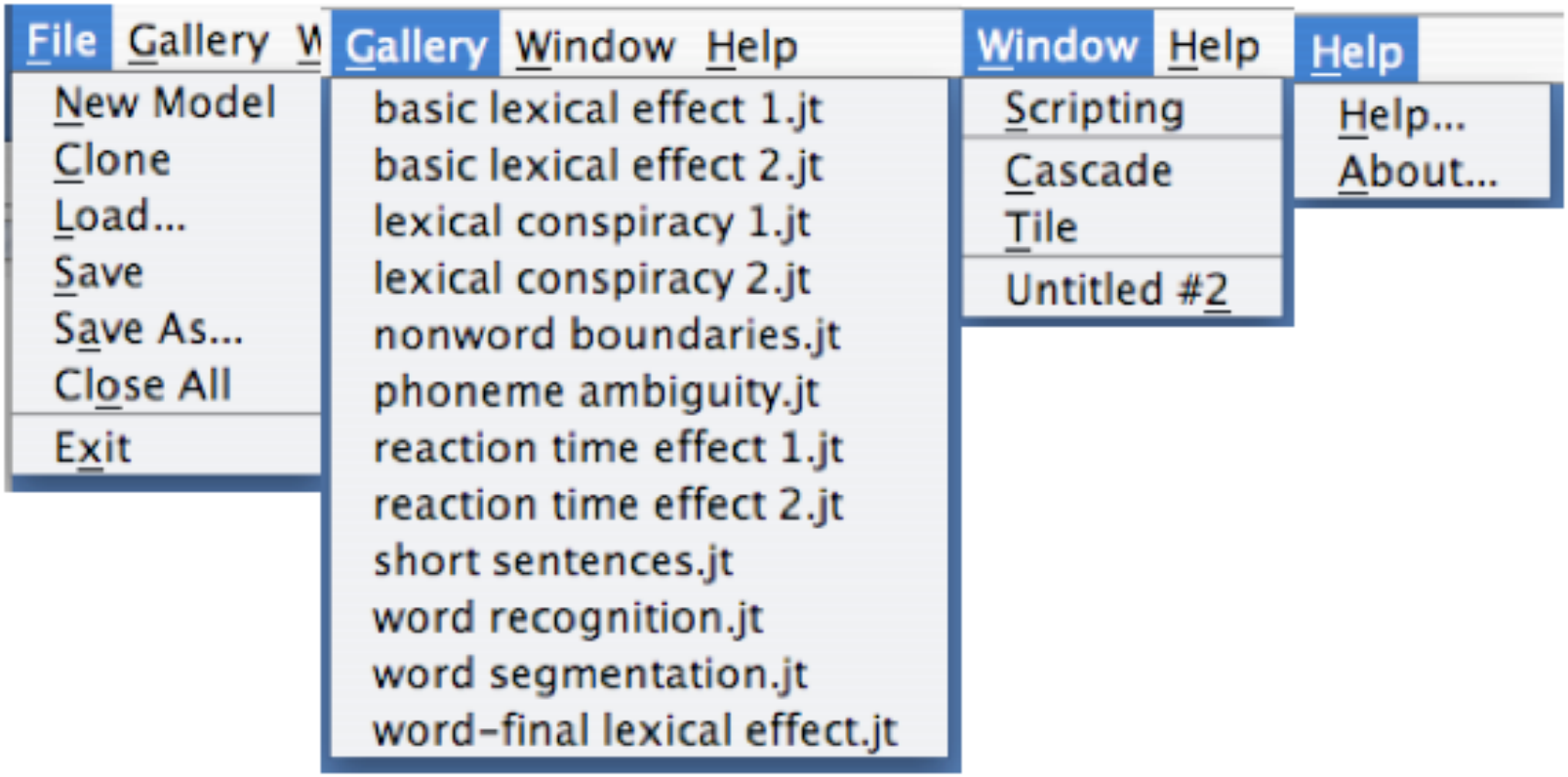
Gallery menu

- Simulations described in the original TRACE paper are implemented and can be executed by simply selecting from this list.
- Simulations and analyses are performed automatically.
- Add to the gallery by saving your simulations to the jtrace/gallery/ folder.



jTRACE Menus

File Gallery Window Help



Your turn

- Try some simulations and explore the interface. Think about what the simulation pertains to perceptual processing.

<u>Example simulations</u>:

1. "-artbist-" : another example of segmentation, note how competitors become inhibited.
2. "-^grit-" : failed segmentation, since *^gri* and *grit* are both in lexicon. How is recognition resolved?
3. "-parti-" : segmentation and competition effects.

Next: Module 3

- 3 classic psycholinguistic experiments, and how their results are modeled by TRACE.
  1. Time course of lexical activation and competition
  2. Lexical consequences of acoustic deviations
  3. Lexical effects on identification of ambiguous phonemes