LEXICAL KNOWLEDGE DOES NOT IMPROVE DISCRIMINABILITY

John Kingston

University of Massachusetts, Amherst
jkingston@linguist.umass.edu

ABSTRACT

The “Ganong effect” [4] refers to listeners’ preference respond to ambiguous steps from a word-nonword continuum with the endpoint category that makes a word. Interactive models such as TRACE [7] attribute this preference to feedback from the lexicon to signal evaluation, while autonomous models such as MERGE [9] attribute it instead to feedforward from the lexicon to a task-specific phoneme-decision module. A categorization and discrimination experiment produced the expected lexical biases in categorization, corresponding shifts in local discrimination peaks, but consistently no greater cumulative discriminability in word-nonword continua than word-word continua. Simulations using jTRACE [10] show that this interactive model neither predicts better discriminability to word-nonword than a word-word continuum nor lexical biases.

Keywords: Interaction, autonomy, discriminability

1. INTRODUCTION

The debate between interactive and autonomous models of speech perception remains unresolved [8, 9]. To contribute new evidence and argument, I assessed the success of simulations using an interactive model, jTRACE [7, 10], in predicting the results of categorization and discrimination experiments.

2. EXPERIMENTS 1A,B

In Exp. 1a, listeners categorized a syllable-final /p-t/ continuum in contexts where either /p/, /t/, or both made a word. In Exp. 1b, they discriminated pairs of stimuli from this continuum.

2.1. Method

2.1.1. Stimuli

Members of a 20-step /p-t/ continua were presented in the contexts /hi-, hu-, ki-, gru-, mi-, ʃu-/, yielding no-(lexical)-bias word-word continua, *heap-heat* and *hoop-hoot*, /p/-bias word-nonword continua, *keep-*keet* and *group-*groot*, and /t/-bias nonword-word continua, **meep-meet* and **shoop-shoot*. The continua were made by manipulating the vowel-to-consonant formant transitions and mixing the /p/ and /t/ bursts in complementary proportions.

2.1.2. Procedures

Exp. 1a: Categorization. Listeners were trained to respond “p” or “t” with correct-answer feedback with 10 repetitions of each continuum endpoint. In the ensuing 6 test blocks, steps {1, 20}, {4, 6, 16, 18}, and {8, 10, 12, 14} were presented in a 1:2:3 ratio. Stimulus presentation was randomized in both training and testing. Each trial comprised a 500-ms display of a cross, the stimulus, a 1500-ms response interval, a 750-ms display of correct-answer feedback on training but not test trials, and a 750-ms ITI.
Exp. 1b: Discrimination. The format was AX, with a 750-ms ISI. In different trials, stimuli differed by 4 steps: 1 vs 5, 3 vs 7, 5 vs 9, 7 vs 11, 9 vs 13, 11 vs 15, 13 vs 17, and 15 vs 19. Stimulus presentation was blocked by stimulus pair. A block began with randomized unscored training trials in which all 4 possible stimulus orders were presented once in each of the six contexts. The 4 possible orders were then presented 6 times each in all 6 contexts, yielding 144 randomized test trials per block, twice on separate days, yielding 24 same and 24 different trials/stimulus pair/context/participant. Order of test blocks was counterbalanced across subjects with a balanced Latin square. Each trial comprised a 500-ms display of a cross, the 2 stimuli, a 1500-ms response interval, a 750-ms display of correct-answer feedback, and a 750-ms ITI.

2.1.3. Listeners

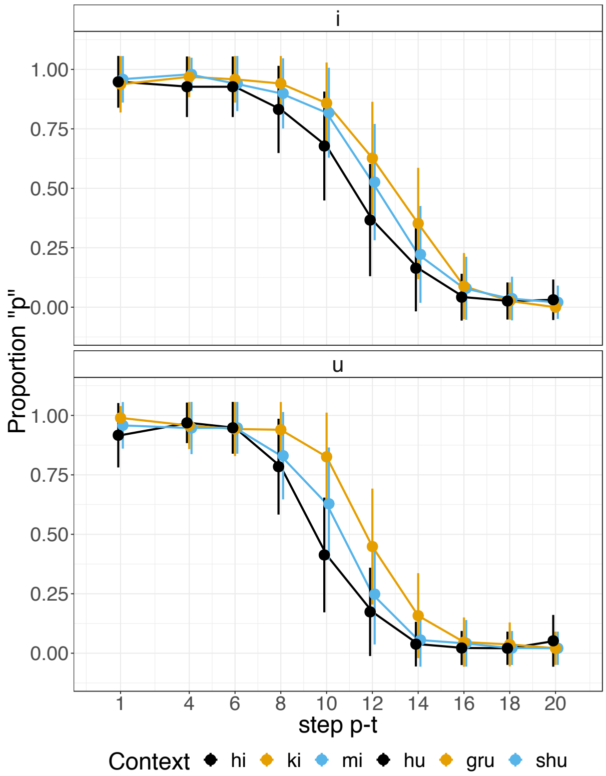
16 adult native monolingual speakers of American English participated in Exp. 1a, and another 33 in Exp. 1b.

2.2. Results

Exp. 1a: Categorization. Fig. 1 shows that listeners responded “p” more often in the /p/-biased contexts, /ki-/, /gru-/ than the /t/-biased contexts /mi-/,

/ʃu-/. They were also biased to respond "p" more often after /i/ than /u/.

Figure 1: Mean proportion "p" (95% CI) in /i/ (top) and /u/ (bottom) contexts.



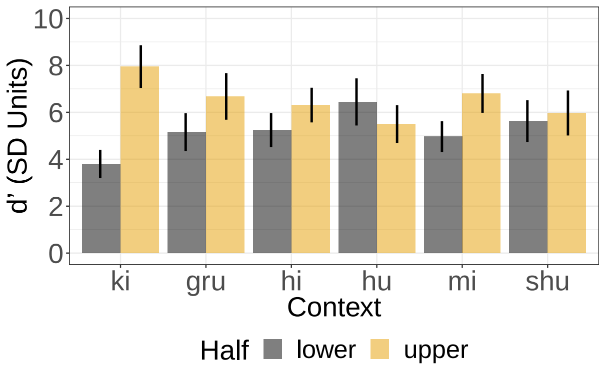
The relative proportion of "p" to "t" responses was submitted to a mixed effects logistic regression model with fixed effects: Step along the /p-t/ continuum, Lexical Bias (/p/-bias: /ki-/, /gru-/ = 0.5; no-bias: /hi-/, /hu-/ = 0; and /t/-bias: /mi-/, /ʃu-/ = -0.5), and Vowel (/p/-bias: /ki-/, /hi-/, /mi-/ = 0.5 vs /t/-bias: /gru-/, /hu-/, /ʃu-/ = -0.5). All fixed effects were scaled. De-correlated random effects of listener on the intercept and the slopes of the fixed effects were included. The estimates in Table 1 show that listeners were not biased to respond "p" (non-significant intercept), they responded "p" less often as the stop became less /p/-like (negative Step), and more often when the lexical context biased them toward "p" (positive Lexical Bias) and when the vowel was /i/ (positive Vowel). No interactions between fixed effects were significant.

Exp. 1b: Discrimination. As expected from Exp. 1a, discriminability (d' values) was greater in the upper half of the continuum nearer to the /t/ endpoint (pairs 9 vs 13 to 15 vs 19) than in the lower half nearer to the /p/ endpoint (pairs 1 vs 5 to 7 vs 11) for the context /hi-/ where the vowel shifted the category boundary away from the /p/ endpoint in Exp. 1a but greater in the lower half for the context /hu-/ where the vowel shifted the boundary toward from /p/ endpoint (Fig. 2). Although discriminability is consistently greater in the upper than the lower half for all the lexically biasing contexts, that difference is still greater for the /p/- than /t/-biased contexts, /ki-/ vs /mi-/ and /gru-/ vs /ʃu-/.

Table 1: Exp. 1a: Fixed effects estimates.

	Est	*se*	*z*	*p*
Intercept	0.231	0.233	-0.989	0.321
Step	-3.299	0.216	-15.271	<0.001
Lexical	0.256	0.052	5.189	<0.001
Vowel	0.417	0.119	3.503	<0.001

Figure 2: Mean lower and upper half cumulative discriminability (95% CI).



Continuum Half (lower = -0.5, upper = 0.5) replaced Step in a mixed effects linear model comparing upper and lower half cumulative discriminability scores (summed d's). De-correlated random effects of listener on the intercept and the slopes of the fixed effects were included. The estimates in Table 2 show that discriminability was greater than chance (positive Intercept) and greater in the upper than the lower half of the continuum (positive Half). Neither Lexical Bias nor Vowel influenced discriminability independently (non-significant main effects), but both interacted significantly with Half (positive interactions with Half): discriminability was relatively greater in the upper than the lower half when either Lexical Bias or Vowel favored "p" in Exp. 1a.

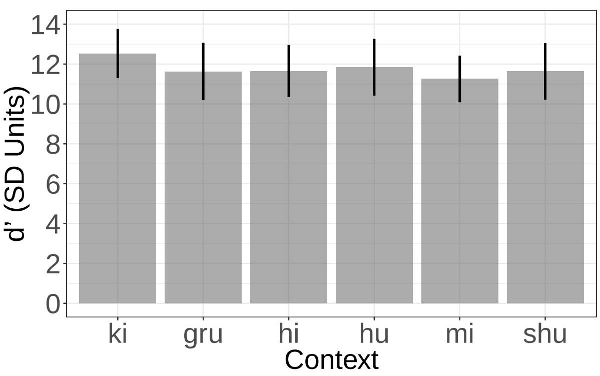
Cumulative discriminability across the continuum (summed d's) differed little between contexts (Fig. 3). A mixed effects linear model produced only a significantly positive intercept. Bayes factors from paired-sample t-tests showed that cumulative d' values from each lexically biasing context do not dif-

Table 2: Exp. 1b: Fixed effects estimates.

	Est	*se*	*t*	*p*
Intercept	5.870	0.296	19.809	<0.001
Half	0.666	0.207	3.225	0.003
Lexical	0.021	0.073	0.293	0.770
Vowel	-0.025	0.081	-0.311	0.758
Half:Lex	0.357	0.074	4.805	<0.001
Half:Vwl	0.512	0.088	5.796	<0.001

fer from those in the corresponding no-bias context: /ki-/ vs /hi-/ 0.719, /mi-/ vs /hi-/ 0.280, /gru-/ vs /hu-/ 0.217, and /ʃu-/ vs /hu-/ 0.227.

Figure 3: Exp. 1b: Cumulative discriminability (95% CI).



2.3. Discussion

These experiments show that the category boundaries and local discrimination peaks shift in concert away from the word endpoint, but that cumulative discriminability across the continuum does not differ between word-nonword and word-word continua.

3. SIMULATIONS

jTRACE simulations were run to test the prediction that feedback from the lexicon improves cumulative discriminability for a word-nonword compared to a word-word continuum, e.g., for *keep-*keet* vs *heap-heat*. This prediction arises if:

1. The /p/ node were excited more by the features of the pair member closer to the /p/ endpoint than by the member further away,
2. The *keep* node were in turn excited more by the closer member,
3. The *keep* node's greater excitation fed back more strongly to the /p/ node for the closer than the further member,
4. No feedback from the lexicon excited the /t/ node more for the further than the closer member,
5. The /p/ node inhibited the /t/ node more for the closer than the further member.

Jointly, these effects would increase /p/ activation and decrease competing /t/ activation more for the closer than the further member in the word-nonword *keep-*keet* continuum than in the corresponding word-word *heap-heat* continuum. Since response probability depends on activation, and discriminability is predicted from differences in response probabilities between members of stimulus pairs, better pair-wise and cumulative discriminability is predicted for a word-nonword continuum than a word-word continuum.

A 202-word lexicon was constructed comprising all the CVC, CV, and VC English words from IPhOD2 [11] that could be generated with jTRACE's phoneme inventory. Homophones were treated as single words and their SubtLex frequencies [1] were pooled. Non-default parameter values were phoneme-word weight 0.05, word-phoneme weight 0.02, feature decay 0.02, word resting activation -0.1, log frequency phoneme-to-word weight 0.13 [2], and stochasticity 0.02 [6]. The possible outputs were the stop consonants /p, t, k, b, d, g/ or a word boundary, the exponent of the Luce forced-choice rule was 7, and alignment was 8 [3]. Each simulation ran for 80 cycles. Each member of a 5-step /p-t/ continuum was simulated 200 times for these contexts (frequencies per million words for words) for: Coda targets: Word-word: *sheep* (13.43)-*sheet* (11.61), *seep* (0.18)-*seat* (78.78), *shop* (53.86)-*shot* (227.43), *soup* (25.2)-*suit* (68.61), *loop* (8.65)-*loot* (3.92); Word-nonword: *deep* (76.39)-**deet*, *leap* (6.67)-**leat*, *sop* (0.27)-**sot*, *dupe* (0.65)-**dute*; Nonword-word: **roop-root* (31.82), **shoop* (0.2)-*shoot* (168.55), **dop-dot* (6.63), **rop-rot* (7.73); Onset targets: Word-word: *peak* (11.52)-*teak* (0.39), *peel* (5.82)-*teal* (0.67), *pool* (46.98)-*tool* (10.75), *pol* (0.45)-*tol* (0.73); Word-nonword: *piece* (194.2)-**tiece*, *peep* (4.43)-**teep*, *poop* (5.59)-**toop*, *pos* (0.31)-**tos*; Nonword-word: **peague-teague* (0.2), **peer-tear* (1.06), **pube-tube* (16.43), **poot-toot* (1.57).

Figs. 4 and 5 show *z*-scores for the median probability of a "p" response between cycles 45 and 75. These medians change equally quickly for word-word, word-nonword, and nonword-word continua. Simulations pooled within continuum type do not produce the expected lexical biases (Table 3).

D_a values were calculated for each pair of stimuli (Eq. 1): H = P("p") to the more /p/-like member of

Table 3: Total proportions “p” (95% CI).

Bias	Coda	Onset
word-word	0.443 (0.080)	0.473 (0.073)
word-nonword	0.457 (0.082)	0.482 (0.073)
nonword-word	0.457 (0.082)	0.473 (0.072)

the pair, F = P(“p”) to the less /p/-like member, s = the ratio of the H and F z-score ranges between their 0.16-0.84 quantiles (Ch. 3 [5]).

(1) $$d_a = \sqrt{\frac{2}{1+s^2}}\,[z(H) - sz(F)]$$

Figure 4: z(P(“p”), simulations (points), codas, loess smoothers: (black) word-word, (ochre) word-nonword, (blue) nonword-word continua.

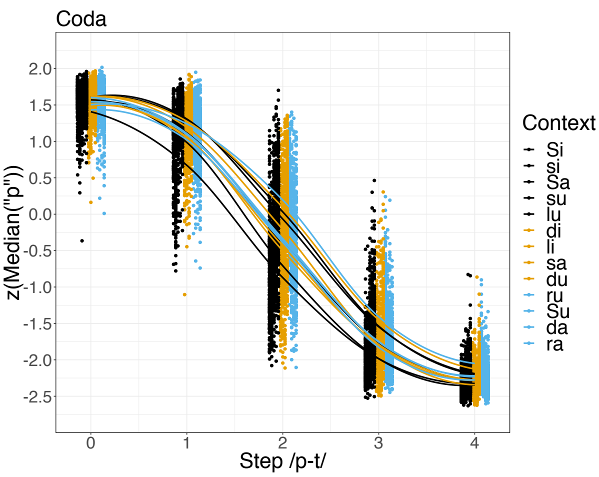
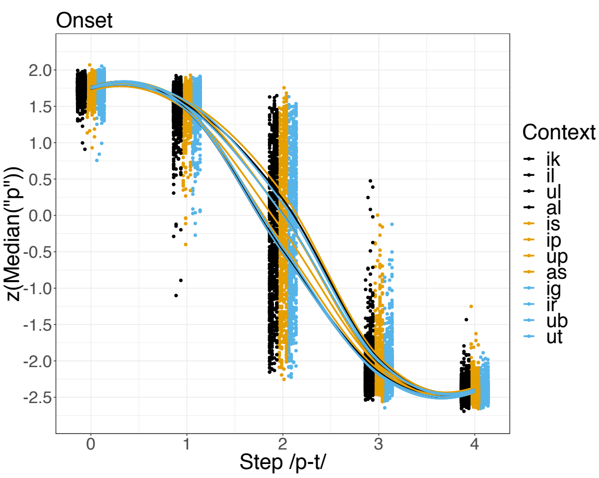


Figure 5: z(P(“p”), onsets.



Figs. 6 and 7 show that simulated stimuli were predicted to be no more discriminable across the continuum for the word-nonword nor nonword-word than word-word continua.

Figure 6: Cumulative d_a. Contexts as in Fig. 4.

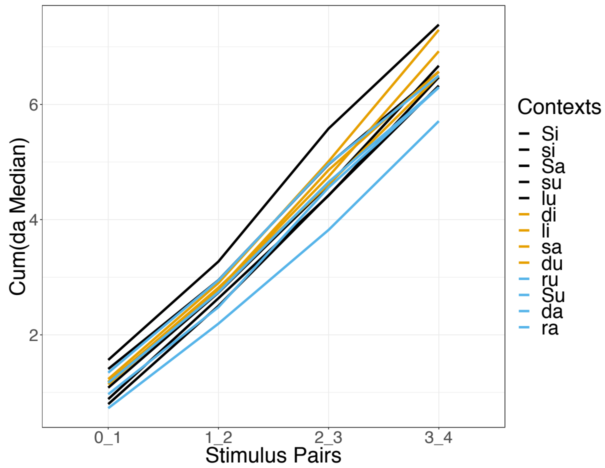
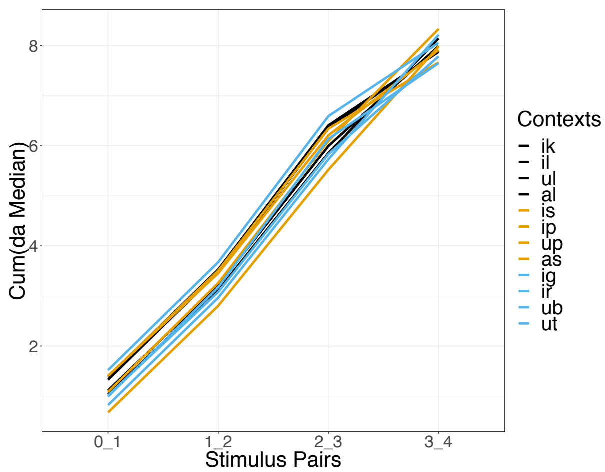


Figure 7: Cumulative d_a. Contexts as in Fig. 5.



4. CONCLUSION

These simulations predict no greater cumulative discriminability for word-nonword nor nonword-word than word-word continua, and therefore that an interactive model would predict the finding in the Exp. 1b that cumulative discriminability would not differ across the three continuum types. But they also don’t produce the expected lexical biases, so do they test the prediction that feedback from the lexicon improves cumulative discriminability for a word-nonword compared to a word-word continuum? (James Magnuson generously helped me in producing the simulations.)

5. REFERENCES

[1] Brysbaert, M., New, B. 2009. Moving beyond Kucera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for american english. *Behavior Research Methods* 41, 977–990.

[2] Dahan, D., Magnuson, J. S., Tanenhaus, M. K. 2002. Time course of frequency effects in spoken word recognition: Evidence from eye movements. *Cognitive Psychology* 42, 317–367.

[3] Frauenfelder, U. H., Peeters, G. 1998. Simulating the time course of spoken word recognition: An analysis of lexical competition in TRACE. In: Grainger, J., Jacobs, A. M., (eds), *Localist Connectionist Approaches to Human Cognition*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers 101–146.

[4] Ganong, W. F. 1980. Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6, 110–125.

[5] Macmillan, N. A., Creelman, C. D. 2005. *Detection Theory: A User's Guide. 2nd Edition*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

[6] McClelland, J. L. 1991. Stochastic interactive processes and the effect of context on perception. *Cognitive Psychology* 23, 1–44.

[7] McClelland, J. L., Elman, J. L. 1986. The TRACE model of speech perception. *Cognitive Psychology* 18, 1–86.

[8] McClelland, J. L., Mirman, D., Holt, L. 2006. Are there interactive processes in speech perception? *TRENDS in Cognitive Sciences* 10, 363–369.

[9] Norris, D., McQueen, J. M., Cutler, A. 2000. Merging information in speech recognition: Feedback is never necessary. *Behavioral and Brain Sciences* 23, 299–370.

[10] Strauss, T. J., Harris, H. D., Magnuson, J. S. 2007. jTRACE: A reimplementation and extension of the TRACE model of speech perception and spoken word recognition. *Behavior Research Methods* 39, 19–30.

[11] Vaden, K. I., Halpin, H. R., Hickok, G. S. 2009. Irvine phonotactic online dictionary, version 2.0.