Phonetic recalibration under concurrent processing of competing representations

Junkai Ma (Junkai.Ma@outlook.com)[1], Brechtje Post[1]
[1]Department of Theoretical and Applied Linguistics, University of Cambridge

Speech perception is assumed to recover discrete phonetic categories from continuous, variable acoustic signals, yet few cues, if any, reliably determine category membership on their own [1, 2]. Comprehension therefore relies on integrating multiple, partially informative cues to estimate the most plausible category, consistent with the view that perception is inference [3, 4]. Perceptual learning keeps inference calibrated when the cue-to-category mapping shifts across talkers and contexts. One such form, phonetic recalibration, shifts the boundary between competing categories when lexical context disambiguates an acoustically ambiguous token, so that the atypical pronunciation is thereafter interpreted as intended [5, 6]. A second form, dimension-based statistical learning, reweights acoustic dimensions by how reliably they distinguish the categories under current conditions, warping perceptual space to raise the signal-to-noise ratio so that reliable cues dominate the decision [7, 8].

Crucially, the cues feeding perceptual inference need not be acoustic. Within a cascading linguistic architecture, representations recovered at each level (phonemes, syllables, words) serve both as processing outcomes and as endogenous cues to inference at other levels [9]. Lexically guided recalibration can thus be reinterpreted as endogenous cue integration, where the recovered lexical-semantic representation identifies the intended category and drives adaptive change [5, 6]. The change itself, however, applies to the cue-to-category mapping rather than the lexical cue that triggers it, so the system must locate where input and expectation diverge, a credit-assignment problem parallel to dimensional reweighting but over representational levels. Adaptation should therefore draw on the resources available at the relevant representational level, allowing concurrent load to probe the representations that recalibration recruits. Dual-task evidence supports this at the grain of modules: concurrent phonological load selectively disrupts adaptation to dysarthric speech (phonologically degraded), whereas no load affects adaptation to noise-vocoded speech (acoustically degraded) [10, 11].

Whether adaptation is similarly organised within the module remains untested. The present study crosses segmental (phonemic) and tonal (syllabic) recalibration with four concurrent loads: visual, segmental, tonal, and lexical. The predictions are twofold: recalibration should be most attenuated by matched-level load, less by load at another phonological level, and least by non-phonological load; any asymmetry in cross-level interference would indicate whether information flows predominantly bottom-up or top-down. Native speakers of Mandarin (N = 48; 12/load group) will first complete load-task familiarisation as a difficulty check, then recalibration under an assigned load (Figure 1). Feedback defines a shifted cue-to-category mapping and drives learning towards it, supplying supervision that overlaps with no load condition. Task order and response mappings will be counterbalanced; post-task ratings will index perceived effort and attention. Trial-level accuracy and RT trajectories will be analysed using Bayesian GAMMs, and ratings using Bayesian hierarchical linear models, all with regularising priors.

References

[1] Liberman, A. M., Cooper, F. S., Shankweiler, D. P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review, 74*(6), 431–461.

[2] Holt, L. L., & Lotto, A. J. (2010). Speech perception as categorization. *Attention, Perception, & Psychophysics, 72*(5), 1218–1227.

[3] Helmholtz, H. von. (1925). *Treatise on physiological optics* (J. P. C. Southall, Trans.; Vol. 3). Optical Society of America. (Original work published 1867)

[4] Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review, 122*(2), 148–203.

[5] Norris, D., McQueen, J. M., & Cutler, A. (2003). Perceptual learning in speech. *Cognitive Psychology, 47*(2), 204–238.

[6] Mitterer, H., Chen, Y., & Zhou, X. (2011). Phonological abstraction in processing lexical-tone variation: Evidence from a learning paradigm. *Cognitive Science, 35*(1), 184–197.

[7] Holt, L. L., Tierney, A. T., Guerra, G., Laffere, A., & Dick, F. (2018). Dimension-selective attention as a possible driver of dynamic, context-dependent re-weighting in speech processing. *Hearing Research, 366*, 50–64.

[8] Goldstone, R. L. (1994). Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General, 123*(2), 178–200.

[9] Martin, A. E. (2016). Language processing as cue integration: Grounding the psychology of language in perception and neurophysiology. *Frontiers in Psychology, 7*, Article 120.

[10] Adank, P., Wang, H., Hepworth, T. J., & Borrie, S. A. (2025). Perceptual adaptation to dysarthric speech is modulated by concurrent phonological processing: A dual task study. *The Journal of the Acoustical Society of America, 157*(3), 1598–1611.

[11] Wang, H., Chen, R., Yan, Y., McGettigan, C., Rosen, S., & Adank, P. (2023). Perceptual learning of noise-vocoded speech under divided attention. *Trends in Hearing, 27*, Article 23312165231192297.

Figure 1 Overview of the experimental procedure.

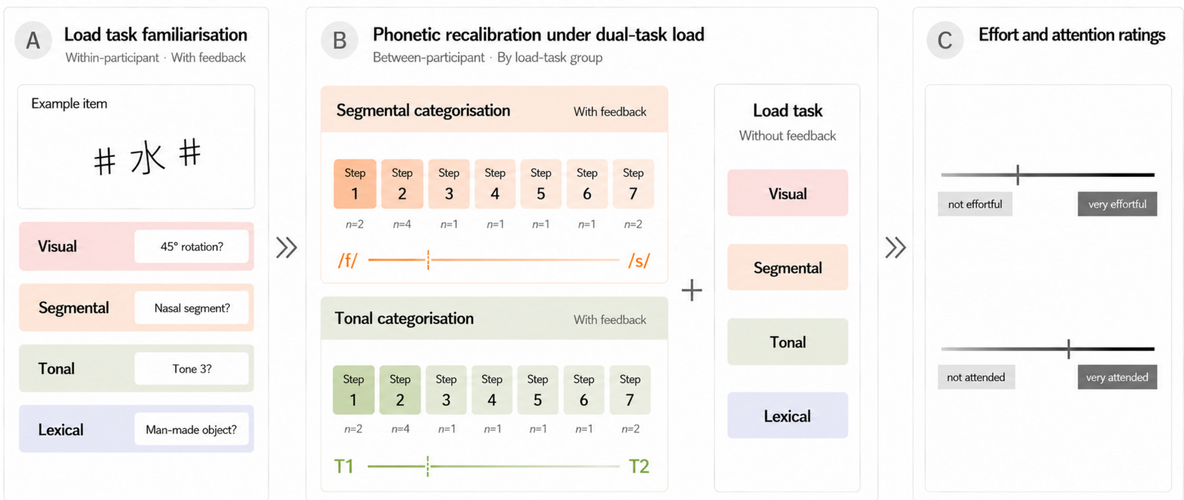

Table 1 Examples of disyllabic pseudowords used in the categorisation task.

Recalibration section	Target contrast	First-syllable endpoint pair		Second-syllable carrier	
		Pinyin	IPA	Pinyin	IPA
Segmental	/f/–/s/	*fen1–sen1*	/fən⁵⁵/–/sən⁵⁵/	*mei2*	/meɪ³⁵/
Segmental	/f/–/s/	*fu1–su1*	/fu⁵⁵/–/su⁵⁵/	*lai2*	/laɪ³⁵/
Tonal	T1–T2	*wa1–wa2*	/wa⁵⁵/–/wa³⁵/	*tie4*	/tʰjɛ⁵¹/
Tonal	T1–T2	*ma1–ma2*	/ma⁵⁵/–/ma³⁵/	*kun4*	/kʰwən⁵¹/

Table 2 Examples of monosyllabic words used in the load task.

Character	Pinyin	Visual orientation	Segmental nasality	Tonal identity	Semantic category	Stroke count
桨 'oar'	*jiang3*	Non-45°	Nasal	T3	Man-made	10
氧 'oxygen'	*yang3*	Non-45°	Nasal	T3	Natural	10
剑 'sword'	*jian4*	45°	Nasal	T4	Man-made	9
象 'elephant'	*xiang4*	45°	Nasal	T4	Natural	11
笔 'pen'	*bi3*	45°	Non-nasal	T3	Man-made	10
海 'sea'	*hai3*	45°	Non-nasal	T3	Natural	10
票 'ticket'	*piao4*	Non-45°	Non-nasal	T4	Man-made	11
树 'tree'	*shu4*	Non-45°	Non-nasal	T4	Natural	9