SWITCHING BETWEEN PHONOLOGICAL BIASES IS NOT FREE: EVIDENCE FROM A MULTILINGUAL RECONSTRUCTION TASK

Adam A. Bramlett, Seth Wiener

Carnegie Mellon University
abramlet@andrew.cmu.edu, sethw1@cmu.edu

ABSTRACT

Language users develop phonological biases in response to the acoustic-phonetic and phono-lexical characteristics of the language being learned. English speakers, for example, develop a consonantal bias (C-bias) whereas Mandarin speakers develop a vocalic bias (V-bias). L1 Mandarin-L2 English multilinguals must switch between an L1 V-bias and an L2 C-bias during L1 and L2 lexical processing. This exploratory study uses the word reconstruction task to examine the potential switch-cost associated with two different phonological biases. Seventeen participants were asked to orally change non-words into words by replacing either the consonant or vowel. Half the participants first did the task in English (C-bias) followed by Mandarin (V-bias) while the other half did the task first in Mandarin (V-bias) followed by English (C-bias). A switch-cost was only found in the Mandarin results in which a V-bias was attenuated. Switching between phonological biases can therefore potentially affect lexical processes in multilinguals.

Keywords: Phonological biases, Switch-cost, Learning, Multilingualism, L2 speech processing

1. INTRODUCTION

All documented spoken languages have consonants and vowels [8]; for this reason, researchers posited that speech sounds must play language-independent roles during language processing [12]. The 'division-of-labor hypothesis' stated that vowels carry a higher functional load in syntactic and prosodic processing, while consonants carry a higher functional load in lexical processing [13]. However, developmental findings tell a more nuanced story, where a C-bias or V-bias is believed to develop in early infancy (e.g., [7], [3], [11], [18]).

For example, [6] used a word learning task to examine French infants' phonological biases. Infants were presented with two objects named by two non-words that involved a one-feature consonantal (e.g., /paʃ-baʃ/) or vocalic contrast (e.g., /ʃyl/-/ʃul/). French infants demonstrated a C-bias: they successfully learned non-words that had consonant contrasts but unsuccessfully learned non-words that had vowel contrasts. In contrast, Danish infants performing a similar task demonstrated a V-bias: they successfully learned words contrasted by a vowel rather than a consonant [7].

For adults, one of the many experimental tasks used to assess phonological biases is the reconstruction task [14]—the task used in the present study. In this task, participants hear a non-word, such as /wɛsk/, and must orally report a real word by changing a single consonant or vowel. For example, by changing the final consonant from /k/ to /t/ for the word /wɛst/, or by changing the vowel from /ɛ/ to /ɪ/ for the word /wɪsk/. Participants are also given a free choice condition in which they can choose either V- or C- substitution. The non-word in this task acts as a "perceptual template" [14] of multiple real words. Participants' correct reaction time (RT) and accuracy serve as the dependent variables. [14]'s English-speaking participants were both faster and more accurate in the vowel change condition (e.g., /wɛsk/ to /wɪsk/) than the consonant change condition (e.g., /wɛsk/ to /wɛst/), and changed vowels more often and faster than consonants in the free change condition. English (and Spanish [2]) listeners are biased towards preserving consonants and thus demonstrate a C-bias.

[21] extended the reconstruction task to Mandarin, thereby adding a fourth condition: tone change. Similar to V- and C- reconstruction, T- reconstruction requires the participant to change a tone to make a non-word into a real word. For example, /taɪ/-*dai2* is a non-word in Mandarin that can be changed into a real word by substituting a vowel (e.g., /tɤ/-*de2*), consonant (e.g., /laɪ/-*lai2*), or tone (e.g., /taɪ/-*dai4*). [21] found a V-bias for Mandarin listeners. That is, Mandarin listeners were least accurate and slowest in vowel reconstruction as compared to tone or consonant reconstruction.

Thus, listeners develop language specific phonological biases. L1 Mandarin [21] and

Cantonese [5] adult listeners, as well as infant L1 Danish [7] listeners, demonstrate a lexically related V-bias whereas L1 English [14], French [6], and Spanish [2] listeners demonstrate a lexically related C-bias. These biases are the result of the acoustic-phonetic and phono-lexical characteristics of the language environment: Danish consonants exhibit a high degree of lenition, which weakens the functional load of consonants [15] whereas Mandarin and Cantonese vowels carry lexically contrastive fundamental frequency information [4]. This tonal information plays a crucial role in lexical access, which increases the functional load of vowels (e.g., [10], [20]).

More recently, [19] examined whether L1 Mandarin listeners (L1 V-bias) demonstrate a new bias in their L2 English reconstruction behavior (i.e., an L2 C-bias). Participants performed the reconstruction task in their L2 and indeed showed faster RTs and lower error rates in the vowel change condition than the consonant change condition. [19] also tested L1 English-L2 Mandarin listeners engaged in classroom L2 learning. These L2 listeners had a reduced lexicon and limited experience with Mandarin; yet, they demonstrated a robust V-bias by changing consonants faster and more accurately than vowels. Lexically related phonological processing biases are not determined solely by a listener's L1, but rather, are at least partially determined by the phono-lexical features of the language.

Whereas the learning of phonological bias happens quickly in the L2, the processing cost of switching between V- and C- bias languages is unknown. In this exploratory study, we examine the effects of switching from a C-bias language to a V-bias language (and vice-versa) using the word reconstruction task. Here we ask two research questions:

1. To what extent is reconstruction accuracy affected by language task order?
2. To what extent is reconstruction reaction time affected by language task order?

We have no a priori reason to believe that doing the task in one language is qualitatively different than doing the task in another language. Naturally, performing the task in the L2 (here, English) may result in slower RTs and lower accuracy scores relative to the group performing the task in their L1. We are interested in comparing performance between subjects who perform the task first in English (C-bias) and second in Mandarin (V-bias) and subjects who perform the task in the opposite order: Mandarin-English. Does a switch-cost exist?

2. METHODS

2.1. Participants

Eighteen L1 Mandarin speakers (3 male; 15 female; mean age = 21.8; SD = 2.0) from Mainland China took part in the study. All participants had normal hearing and speech, had completed up to high school in China, and were studying at an American university. All participants reported Mandarin as the only Chinese language spoken or understood. One participant's data was removed for failing to follow the task's instructions.

2.2. Materials

For the English stimuli, 60 non-words were taken from [14]. Each non-word could be changed into a real word by either changing a vowel or a consonant. The position of the substitution was controlled with 35 consonant changes happening before vowel changes (e.g., /wɛmən/ to /lɛmən/ or /wʊmən/) and 25 vowel changes occurring before consonant changes (e.g., /wɛsk/ to /wɪsk/ or /wɛst/). An additional 70 non-words were created as fillers along with 12 practice items. For the Mandarin stimuli, 64 CV and CVV non-words were taken from [21]. Each non-word could be made into a real word by changing either consonant, vowel, or tone. In contrast to the English stimuli, the Mandarin stimuli always had consonant changes before the possible vowel changes. This is due to the limited syllable structure of Mandarin words [19]. 32 CVN nasal final words were used as fillers and 16 more were used as practice items. All stimuli were recorded at 44.1 kHz and 16-bit resolution in a sound attenuated booth. All stimuli, data, and R code are available on the Open Science Framework: https://osf.io/9zgd3/.

2.3. Procedure

Participants were tested individually in a lab setting using headphones. First, participants answered a language background questionnaire, then were given printed and oral instructions in the target language. Order of language was counterbalanced: nine performed the English version then the Mandarin version; nine performed the Mandarin then the English version. Within each language, order of the change conditions (e.g., C-, V-) was counterbalanced. In each change condition, participants were asked to change a speech sound so that a real word could be created. Participants were given 4 practice trials with feedback. Participants were told to focus on the word's sound and not the spelling, and to say the word into a microphone as quickly and clearly as possible. Stimuli were presented using E-Prime [16]. Each trial was separated by a button click controlled by the participant with a 2 second ISI and a 10 second time limit per trial. Verbal responses and reaction times were recorded at word onset using Chronos [1]. Participants were unaware that they would be doing multiple conditions and the task in different languages. Oral responses were transcribed by two bilingual Mandarin-English speakers.

2.4. Results

~10.78% of all responses were removed due to timing out at 10 seconds (Mandarin: ~9.07%, English: ~12.31%). All single changes that were appropriate for the respective task were marked as correct. Responses that had more than one change (e.g., /wɛsk/ to /tɛst/) or had an incorrect change for the respective condition were marked incorrect. Three English words were removed from analysis for low accuracy (/mɪtən/, /lɒvəl/, /gruːn/). Median absolute deviation was used for removal of log transformed reaction times [9] removing 47 items. Mean reaction times of correct responses and error rates of each task can be found in Table 1. Table 1 shows that in the Mandarin task, on average, vowel changes were the slowest and had the highest error rate (i.e., listeners showed a V-bias); whereas, in the English task, on average, consonant changes were the slowest and had the highest error rate (i.e., listeners showed a C-bias).

Task Language	Reconstruction Condition	Mean RT(ms)	Mean % error
Mandarin	Vowel	4934.5	35.37
Mandarin	Tone	2562.1	19.23
Mandarin	Free	2888.0	19.14
Mandarin	Consonant	3637.0	29.04
English	Vowel	3237.0	30.49
English	Free	3029.1	34.36
English	Consonant	3348.7	45.90

Table 1: Mean correct reaction times (ms) and percent error for each task.

To analyze switch-cost, correct reaction time and accuracy were compared across language testing orders for English and Mandarin. To do this, we created a variable for the order in which the task was done (i.e., whether the English/Mandarin task was done first or second). For accuracy, (in)correct answers were analyzed as a binary variable. For reaction time, only correct reaction times were analyzed, following [14].

To test whether there is a switch-cost associated with doing the task in English (C-bias) or Mandarin (V-bias) and vice-versa, two generalized mixed effects models (GMEM) were built to compare accuracy and two linear mixed effects models (LMEM) were used to compare log transformed reaction time results in R [17]. Models were built using the *lme4* package. The models included two fixed effects: task order (first or second), condition (vowel, tone, consonant, free) and their interaction. Task order (first) and condition-consonant were set as reference levels. All variables were coded with effects coding. Random intercepts for item and participant were included, as were random slopes for conditions but these slopes were removed after non-convergence. An effect of condition would indicate a C- or V-bias whereas an effect of task order and/or an interaction would indicate a switch-cost.

Accuracy results from the English task GMEM found null results at an alpha-level of .05. Accuracy results from the Mandarin task GMEM found a significant effect in the free (β = 1.00, SE = .42, z = 2.35, p < .05), tone (β = .99, SE = .43, z = 2.30, p < .05), and vowel conditions (β = -1.44, SE = .41, z = -3.45, p < .001), indicating more accurate reconstruction for free and tone conditions and less accurate vowel reconstruction in comparison to consonants. Additionally, a significant interaction was found between the vowel condition and task order (β = 1.10, SE = .47, z = 2.33, p < .05), indicating that there is a switch-cost, which results in more accurate responses when the Mandarin vowel task is performed after the English tasks.

RT results from the English task LMEM found null effects. The Mandarin RT LMEM found effects for the free condition (β = -.35, SE = .07, t = 5.17, p < .001), tone condition (β = -.57, SE = .69, t = -8.21, p < .001), and vowel condition (β = .79, SE = .08, t = 10.56, p < .001), indicating that vowel responses had significantly slower RTs while tone and free responses had significantly faster RTs (relative to consonant responses). Additionally, an interaction between task order and the free condition was found (β = -.18, SE = .09, t = -2.17, p < .05), indicating that doing the free condition in Mandarin second allowed for faster reaction times.

3. DISCUSSION

Like the results of [19] and [21], our Mandarin reconstruction results offer additional evidence of Mandarin speakers demonstrating a V-bias: L1 Mandarin speakers responded slower and less accurately towards vowel reconstruction than consonant reconstruction. Unexpectedly, the C-bias that was reported in [14] was not replicated. Numerically, however, English consonant accuracy (and RT) was lower than vowel accuracy (and slower than vowel RT). Keeping in mind that our participants were doing the English task in their L2, their means (Table 1) are roughly comparable with [14]'s L1 English participants' means (vowel: 2,217 ms, 28%; consonant: 2,412 ms, 42%). The error rates are nearly identical while our L2 participants were about a second slower, on average. The null effect in our English task is most likely due to the low power of this study (17 participants). Additionally, 15% of the variability found in the English accuracy model was due to participant variability. We are in the process of resuming in-person testing in order to obtain a larger sample.

With respect to switch-cost, we found interactions in the Mandarin accuracy and RT models. When performing the Mandarin reconstruction task after the English reconstruction task participants became more accurate at vowel changes and faster at free choice changes. Figure 1 shows this switch-cost in terms of the interaction between task order and vowel condition accuracy. Figure 2 shows this switch-cost in terms of the interaction between task order and free condition RT. These figures show model outputs by displaying the standard error with horizontal blocks and a vertical center line which marks coefficient estimates. Teal and yellow bars represent coefficient estimates that are lower (teal) and higher (yellow) than reference levels (i.e., condition: consonant, task order: Mandarin-first), respectively. Significant results are marked with red asterisks ($p < .05$ = *, $p < .001$ = ***). In general, doing a task second appears to lower reaction time in the free choice condition (a practice effect), which is demonstrated by the significant effects found in the Mandarin models (Fig 2) and numerical difference in the English task. In contrast, the significant interaction between vowel condition and task order in accuracy suggests that the V-bias demonstrated in Mandarin tasks is somewhat attenuated when a participant first does the English tasks (Fig 1). Doing the English task second leads to a numerically greater error in the vowel change condition, though this effect is not significant at an alpha-level of .05.

Additionally, eight post-hoc analyses were performed by task order and language for both accuracy and RT. In Mandarin reconstruction, vowels were significantly slower despite task order (i.e., V-bias persisted across task order). In contrast, vowels were only significantly less accurate

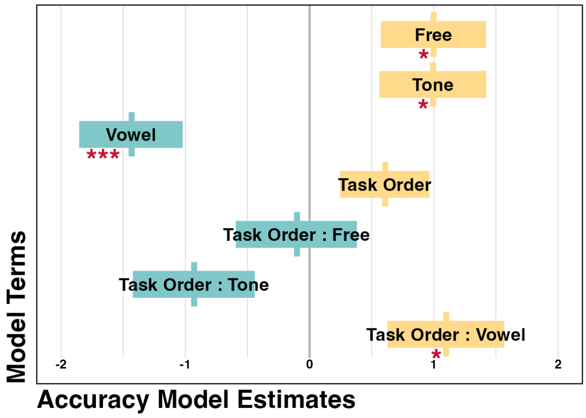

Figure 1: Mandarin accuracy model output

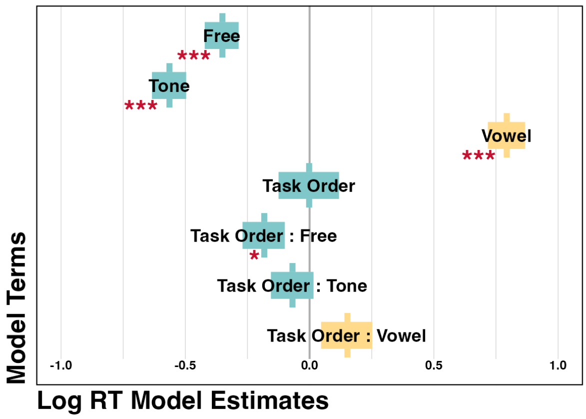

Figure 2: Mandarin Log RT model output

when doing Mandarin first. For English-first participants, however, no V-bias was found, which may indicate an increase in flexibility for lexical retrieval. English RT post-hoc analyses showed null effects. Accuracy results in English-first condition, however, indicated that vowels are more accurate than consonants (i.e., showing a C-bias; [14]). The increased accuracy for vowels in the English-first group is lost in the post-hoc analysis for English-second where no significant difference is found. This change indicates a potential loss of flexibility for vowel-driven lexical access for those that do Mandarin first then English.

In sum, this exploratory study demonstrated a processing cost for multilinguals switching between phonological biases. Listeners demonstrate language-specific patterns of biases, and yet these patterns can change given the previously processed language. To what degree these findings impact theories of multilingual lexical representations and processing, and whether these findings can improve adult language learning remain open questions for future research to address.

REFERENCES

[1] D. L. Babjack, B. Cernicky, A. J. Sobotka, *et al.*, "Reducing audio stimulus presentation latencies across studies, laboratories, and hardware and operating system configurations," *Behav. Res. Methods*, vol. 47, pp. 649–665, 2015.

[2] A. Culter, N. Sebastián-Gallés, O. Soler-Vilageliu, and B. Van Ooijen, "Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons," *Mem. Cognit.*, vol. 28, no. 5, pp. 746–755, 2000.

[3] C. Floccia, T. Nazzi, C. Delle Luche, S. Poltrock, and J. Goslin, "English-learning one- to two-year-olds do not show a consonant bias in word learning," *J. Child Lang.*, vol. 41, no. 5, pp. 1085–1114, 2014.

[4] J. Gandour, "Tone perception in far eastern-languages," *J. Phon.*, vol. 11, no. 2, pp. 149–175, 1983.

[5] D. M. Gómez, P. Mok, M. Ordin, J. Mehler, and M. Nespor, "Statistical speech segmentation in tone languages: The role of lexical tones," *Lang. Speech*, vol. 61, no. 1, pp. 84–96, 2018.

[6] M. Havy and T. Nazzi, "Better processing of consonantal over vocalic information in word learning at 16 months of age," *Infancy*, vol. 14, no. 4, pp. 439–456, 2009.

[7] A. Højen and T. Nazzi, "Vowel bias in danish word-learning: Processing biases are language-specific," en, *Dev. Sci.*, vol. 19, no. 1, pp. 41–49, 2016.

[8] P. Ladefoged and S. F. Disner, *Vowels and Consonants*, 3rd ed. Chichester, England: Wiley-Blackwell, 2012.

[9] C. Leys, C. Ley, O. Klein, P. Bernard, and L. Licata, "Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median," *J. Exp. Soc. Psychol.*, vol. 49, no. 4, pp. 764–766, 2013.

[10] J. G. Malins and M. F. Joanisse, "The roles of tonal and segmental information in mandarin spoken word recognition: An eyetracking study," *J. Mem. Lang.*, vol. 62, no. 4, pp. 407–420, 2010.

[11] T. Nazzi and S. Poltrock, "Consonant bias in the use of phonological information during lexical processing: A lifespan and cross-linguistic perspective," in *Speech perception and spoken word recognition*, J. Gaskell, Ed., London, England: Psychology Press, 2016, pp. 37–54.

[12] T. Nazzi and A. Cutler, "How consonants and vowels shape spoken-language recognition," *Annu. Rev. Linguist.*, vol. 5, no. 1, pp. 25–47, 2019.

[13] M. Nespor, M. Peña, and J. Mehler, "On the different roles of vowels and consonants in speech processing and language acquisition," *Lingue e linguaggio*, vol. 2, pp. 203–230, 2003.

[14] B. van Ooijen, "Vowel mutability and lexical selection in english: Evidence from a word reconstruction task," *Mem. Cognit.*, vol. 24, no. 5, pp. 573–583, 1996.

[15] N. Pharao, "Plosive reduction at the group level and in the individual speaker," in *ICPhS*, XVII, Hong Kong, 2011.

[16] Psychology Software Tools, *E-Prime*, Pittsburgh, PA.

[17] R Core Team, *R: A Language and Environment for Statistical Computing*, Vienna, Austria, 2022. [Online]. Available: https://www.r-project.org/.

[18] L. Singh, H. H. Goh, and T. D. Wewalaarachchi, "Spoken word recognition in early childhood: Comparative effects of vowel, consonant and lexical tone variation," *Cognition*, vol. 142, pp. 1–11, 2015.

[19] S. Wiener, "Second language learners develop non-native lexical processing biases," *Bilingualism*, vol. 23, no. 1, pp. 119–130, 2020.

[20] S. Wiener and K. Ito, "Do syllable-specific tonal probabilities guide lexical access? evidence from mandarin, shanghai and cantonese speakers," *Lang. Cogn. Neurosci.*, vol. 30, no. 9, pp. 1048–1060, 2015.

[21] S. Wiener and R. Turnbull, "Constraints of tones, vowels and consonants on lexical selection in mandarin chinese," *Lang. Speech*, vol. 59, pp. 59–82, 2016.