



# Native phonotactic interference in L2 vowel processing: Mouse-tracking reveals cognitive conflicts during identification

Yizhou Wang<sup>1</sup>, Rikke L. Bundgaard-Nielsen<sup>2</sup>, Brett J. Baker<sup>1</sup>, Olga Maxwell<sup>1</sup>

<sup>1</sup>School of Languages and Linguistics, University of Melbourne, Australia

<sup>2</sup>MARCS Institute for Brain, Behaviour and Development, University of Western Sydney, Australia

yizwang3@unimelb.edu.au, rikkelou@gmail.com, bjbaker@unimelb.edu.au

omaxwell@unimelb.edu.au

## Abstract

Regularities of phoneme distribution in a listener's native language (L1), i.e., L1 phonotactics, can at times induce interference in their perception of second language (L2) phonemes and phonemic strings. This paper presents a study examining phonological interference experienced by L1 Mandarin listeners in identifying the English /i/ vowel in three consonantal contexts /p, f, w/, which have different distributional patterns in Mandarin phonology: /pi/ is a licit sequence in Mandarin, \*/fi/ is illicit due to co-occurrence restrictions, and \*/wi/ is illicit due to Mandarin contextual allophony. L1 Mandarin listeners completed two versions of an identification experiment (keystroke and mouse-tracking), in which they identified vowels in different consonantal contexts. Analysis of error rates, response times, and hand motions in the tasks suggests that L1 co-occurrence restriction and contextual allophony induce different levels of phonological interference in L2 vowel perception compared to the licit control condition. In support of the dynamic theory of linguistic cognition, our results indicate that illicit phonotactic contexts can lead to more identification errors, longer decision processes, and spurious activation of a distractor category.

**Index Terms:** vowel perception, phonotactics, Mandarin

## 1. Introduction

Previous research has shown that non-native segments can be challenging to perceive if they are not contrastive phonemes in the listener's native phonology [1], [2]. Studies also suggest that some non-native segments may cause perceptual difficulties in specific phonological contexts, typically forming phonemic sequences that are not licit strings according to the listener's native phonotactics [3]–[5]. In the present study, we test the effect of phonotactic context on L1 Mandarin listeners' perception of English /ei/ and /i/ vowels from a perspective where speech perception is regarded as a dynamic process of information integration [6], [7]. Specifically, the dynamic theory of linguistic cognition contends that spoken word recognition is not strictly stage-based (i.e., the neural subsystems wait until a stable representation is computed before the information is passed on to the next processing stage), but rather, speech perception involves multiple parallel processes for evidence extraction and accumulation, and over time information from different pathways converge onto a stable and integrated response. Such hypotheses could be tested with technologies which tap into online processing of linguistic stimuli in real time, e.g., eye-tracking [8] and mouse-tracking [7].

In this paper, we aim to explore the behavioural correlates of perceptual interference during L2 vowel identification in familiar and unfamiliar phonotactic contexts using two versions (keystroke and mouse-tracking) of the same identification task. We examine whether non-native listeners experience spurious activation of distractor categories in unfamiliar phonotactic contexts, which motivates potential phonological repairs, i.e., where the target vowel is misidentified in accordance with native phonotactic constraints. Methodologically, we also examine the effectiveness of the mouse-tracking paradigm [7], [9] as compared to the more traditional keystroke paradigm. In a keystroke paradigm, error rate (%Error) and response time (RT) are standard measures for evaluating perceptual difficulty in L2 listeners. While these are excellent metrics for revealing the difficulty level and the degree of effort expended, they are unable to reveal the details of the decision-making process. Therefore, the mouse-tracking technique may complement the identification procedure with extra information about the underlying cognitive process, e.g., the time course of conflict emergence and resolution, which can be analysed from the curvature (spatial) and dynamic (temporal) patterns in the mouse trajectories [7], [9]–[11].

Accordingly, manipulating task complexity can result in different levels of competition between a distractor category and the target category in a two-alternative task. At the behavioural level, spurious activation of a distractor can lead to more curved mouse trajectories, while trajectories tend to be straight lines when interference is minimal [11]. This has been shown in a study which reported that native English speakers showed more curved mouse trajectories for identifying the object 'candle' while the distractor was a phonological neighbour, e.g., 'candy', as compared to an irrelevant category, e.g., 'jacket' [7]. Hand movement is a valid index of real-time cognitive dynamics because goal-directed motor movement is indicative of response competition, and the relation between speech perception, vision, and action is tightly coupled [12].

Mandarin and English jointly provide an ideal phonological setting for testing the effect of interest in the current study. In Mandarin phonology [13], [14], the maximal permissible syllable structure is CGVN or CGVG (C = consonant, G = glide, V = vowel, N = nasal). Monophthongs are allophonically long in open syllables, and diphthongs are often analysed as VG sequences. Not all onset-rime combinations are attested in the Mandarin lexicon. For example, the labials /f, w/ cannot combine with /i/ to render a licit syllable, and thus English sequences \*/fi/ and \*/wi/ are non-existent (i.e., illicit) in Mandarin at the underlying phonological level [13], [15]. In contrast, the combination /p/ + /i/ is a licit syllable. When the nucleus consists of the diphthong /ei/, all three combinations /pei, fei, wei/ are licit syllables in Mandarin, see Figure 1 for

the interaction between underlying and surface representations, as well as some sample words in Mandarin. Moreover, Mandarin has an optional *Triphthong Raising* process which raises the height of segments between two [high] segments, i.e., [high][mid][high] becomes [high][high][high] [13], [15]. Because of this process, a sequence like /wei/ can be realized as [wi] phonetically. Therefore, although English \*/wi/ is treated as an illicit string underlyingly, its phonetic realisation [wi] should have non-zero frequency in Mandarin.

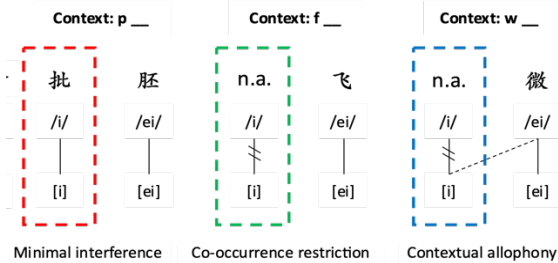


Figure 1. Phonotactic status of CV combinations.

These alternation patterns raise interesting questions for speech perception. We predict that Mandarin listeners will have no difficulty discriminating /pi/ and /pei/ since both sequences are licit, and contrastive, syllables. As for \*/fi/-/fei/, which create an illicit-licit pair in Mandarin phonology, listeners may find it difficult to identify /i/ in contexts of /f/ because Mandarin co-occurrence restrictions prohibit such a combination. Previous research suggests that listeners typically avoid segmental sequences that cause phonotactic violations [3]. Accordingly, we predict that Mandarin listeners will experience some top-down interference due to implicit knowledge of Mandarin phonology and lexicon. If category substitution is a viable strategy for perceptual modification, they may perceive \*/fi/ as /fei/ and experience perceptual confusion. Lastly, \*/wi/-/wei/ creates a slightly more challenging scenario for Mandarin speakers, because speech perception involves reverse inferences, i.e., listeners make the best estimate of the underlying phonological representation by integrating bottom-up sensory information and top-down phonological knowledge [16], [17]. As the *Triphthong Raising* process can cause /wei/ to be realised as [wi] (but not /fei/ as [fi]), Mandarin listeners could interpret a phonetic signal of [wi] as a realization of /wei/. In other words, the contrast between /i/ and /ei/ is neutralised following /w/, and they can be considered contextual allophones. As a result, Mandarin listeners may experience extra interference due to the alternation pattern at the surface level, in addition to the effect of the co-occurrence restriction at the underlying phonological level.

## 2. Methods

### 2.1. Participants

Participants were fifteen native Standard Mandarin speakers (fourteen females and one male,  $M_{age} = 23.5$ ,  $SD = 1.6$ , range: 21-26) with no reported hearing disorders. All were postgraduate students of applied linguistics and spoke English as a second (foreign) language. On average, their age at the onset of English acquisition (AoA) was 7.3 years old ( $SD = 2.2$ , range: 4-12), and they had received classroom-based English education for 14.1 years ( $SD = 3.3$ , range: 7-20) at the time of testing. The participants self-identified as advanced level (B2-

C1) English learners. In addition to Standard Mandarin, some participants spoke regional Mandarin dialects ( $N = 1$  for Henan, Tianjin, Hubei, and Wuhan dialect), but none spoke a third language. The participants were undertaking distance learning, and none had lived in an English-speaking country. The participants gave written consent to participate in the present study and received a small payment for their time.

### 2.2. Stimuli

The stimuli were produced by two phonetically trained male Australian English (AusE) speakers. The stimuli consisted of six CVCV pseudowords /piba/, \*/fiba/, \*/wiba/, /peiba/, /feiba/, and /weiba/, where /i/ represents the tense vowel [i:]. The second syllable /ba/ was used to generate a controlled phonological context, and the labial consonant was chosen to avoid additional co-articulation effects on the neighbouring vowels. The speakers produced the CVCV pseudowords with an AusE stress pattern (the first syllable was stressed), and they produced the stimuli in a clear speech style to maximise the acoustic differences and avoid hypoarticulation [18].

### 2.3. Procedure

The participants completed two versions of an identification task on two consecutive days. The first experiment was a keystroke identification task developed in PsyToolkit 3.3 [19], [20]. The task had 120 trials (two vowels, three consonantal contexts, five tokens for each category, four repetitions) presented in a pseudo-randomized order. Participants were instructed to identify the quality of the first vowel in the /CVba/ stimuli. A stimulus was first played on each trial, and then the two vowel labels (/ei/ or /i/) were printed on the two sides of the screen. Participants pressed the “F” key to identify the category label printed on the left side of the screen or the “J” key to choose the category on the right side. The position of the correct answers was counterbalanced so that potential biases such as the influence of handedness were cancelled out. Participants were encouraged to respond promptly within a time window of 3.0 seconds. See Figure 2 for schematic illustrations of the experiment interface.

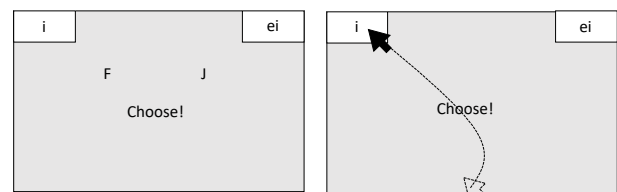


Figure 2. Schematic task interface of the keystroke task (left) and the mouse-tracking task (right).

The second experiment was a mouse-tracking identification task developed in PsychoPy 3.0 [21], [22]. The mouse-tracking identification task had the same condition setting as the keystroke task, but participants made identification responses by mouse clicking. One participant was left-handed while they used a right-handed mouse on a daily basis, and they completed the task also with the right hand. Participants were tested one day after they completed the keystroke task, and one participant did not complete this task. The screen was normalized to a 2.0 units  $\times$  2.0 units canvas. On each trial, participants clicked a “start” box located in the centre of the screen bottom [0, 0] to play the stimulus sound, during which a cross appeared at the

centre of the screen. After stimulus presentation, two vowel labels showed in the top-left box  $[-1.0, 2.0]$  and the top-right box  $[1.0, 2.0]$ . Participants were instructed to give identification responses by moving their mouse from the starting position to click the labelled choice box. The mouse trajectory between the starting and the ending points was automatically recorded. After a trial, the “start” box was presented again, and participants clicked it to start the next trial from the same initial position. In particular, the sampling rate of mouse trajectory recording was determined by the frame rate (frame per second [FPS]) of their computer screen. On average, the testing frame rate was 59.6 FPS (range: 58.0-60.2). This indicates that on average, the distance between two adjacent mouse positions reflects the on-screen displacement within 16.8 ms. The time courses were normalised into 100 timesteps for each trial. For analysis, all rightward movements were mirrored as leftward trajectories so that the top-left box represented the target category /i/, while the top-right box represented the distractor /ei/.

### 3. Results

#### 3.1. Keystroke identification results

The participants’ %Error and RT metrics in the keystroke identification task are summarised in Table 1. The critical (illicit) condition \*/wi/ has led to the highest % Error ( $M = 13.2$ ,  $SD = 17.3$ ), while the other five conditions had a mean % Error lower than 4%. The lowest % Error was found in the control condition /pi/ ( $M = 0.3$ ,  $SD = 1.3$ ). A Friedman’s ANOVA test revealed a significant difference in % Error,  $\chi^2 = 21.559$ ,  $df = 5$ ,  $p = .0006$ . A series of *post hoc* Conover-Bonferroni rank-sum tests revealed that the %Error of \*/wi/ was higher than all other five conditions ( $p < .0001$  for five comparisons), but there were no statistical differences between the other five conditions ( $p > .18$  for ten comparisons).

Table 1. Error rate and RT in the keystroke task.

CV	Condition	%Error	SD	RT (ms)	SD
/pi/	Licit	0.3	1.3	616	236
*/fi/	Illicit	3.7	7.2	678	277
*/wi/	Illicit	13.2	17.3	777	394
/pei/	Distractor	0.7	1.8	694	340
/fei/	Distractor	2.3	4.6	664	318
/wei/	Distractor	2.7	5.3	633	332

RT data were analysed using a generalised linear mixed-effects model (GLMM, Gamma link) and a Wald Chi-squared test, which showed a significant effect of CV combination on RT,  $\chi^2 = 69.099$ ,  $df = 5$ ,  $p < .0001$ . We then carried out a series of *post hoc* tests based on estimated marginal means (EMMs) with *p*-values adjusted using Bonferroni’s method for fifteen comparisons. The tests revealed that the mean RT of \*/wi/ was significantly longer than /pi/ ( $M_{diff} = 152$  ms,  $p < .0001$ ) and \*/fi/ ( $M_{diff} = 88$  ms,  $p = .0038$ ). In addition, the mean RT of \*/fi/ was significantly longer than /pi/ ( $M_{diff} = 64$  ms,  $p = .0122$ ). There were no significant differences in RT between the three distractor categories ( $p > .14$  for three comparisons).

#### 3.2. Mouse-tracking identification results

The %Error and RT in the mouse-tracking identification task are summarised in Table 2. A Friedman’s rank-sum repeated-measures ANOVA revealed a significant difference in %Error,  $\chi^2 = 19.884$ ,  $df = 5$ ,  $p = .0013$ . *Post hoc* Conover-Bonferroni

tests revealed that the % Error of \*/wi/ was greater than all other five conditions ( $p < .0004$  for five comparisons), but there were no statistical differences between the other five conditions ( $p > .9999$  for ten comparisons). The RT data were again analysed using a GLMM and a Wald test, which revealed no significant differences between conditions in general,  $\chi^2 = 9.9246$ ,  $df = 5$ ,  $p = .0774$ .

Table 2. Error rate and RT in mouse-tracking task.

CV	Condition	%Error	SD	RT (ms)	SD
/pi/	Licit	0	0	1015	289
*/fi/	Illicit	1.1	2.9	996	294
*/wi/	Illicit	7.5	13.0	1073	308
/pei/	Distractor	0	0	1001	289
/fei/	Distractor	0.4	1.3	1049	383
/wei/	Distractor	0.4	1.3	1006	279

The average mouse trajectories for the correct responses are summarised in Figure 3. To analyse the temporal course of the cursor movement, we built a generalised additive mixed-effects model (GAMM) [23] to model the horizontal deviation along the *x*-axis over time (because the *y*-axis was controlled for all conditions). The CV condition was set as a fixed factor, and the participant was set as a random effect. To compare the between-condition differences in time course, we carried out a series of pairwise contrast analyses based on the GAMM, see Figure 4. This revealed that the trajectory of \*/fi/ differed significantly from /pi/ between the 58th and 82ed time steps ( $p < .05$ ), and the maximal difference was 0.073 units at the 74th time step ( $\approx 757$  ms,  $M_{RT} = 1023$  ms). This positive component of the *x*-coordinate indicates the attraction towards the distractor, which was set on the top-right corner  $[1.0, 2.0]$ . The *x*-trajectory of \*/wi/ also differed significantly from /pi/ between the 41st and 74th time steps ( $p < .05$ ), and the maximal difference was 0.117 units at the 55th time step ( $\approx 562$  ms). Lastly, the trajectory of \*/wi/ differed significantly from \*/fi/ between the 44th and 53rd time steps ( $p < .05$ ), and the maximal difference was 0.066 units at the 46th time step ( $\approx 470$  ms). We also compared the *x*-trajectories of the distractors /pei/, /fei/, and /wei/, but those three trajectories did not differ significantly.

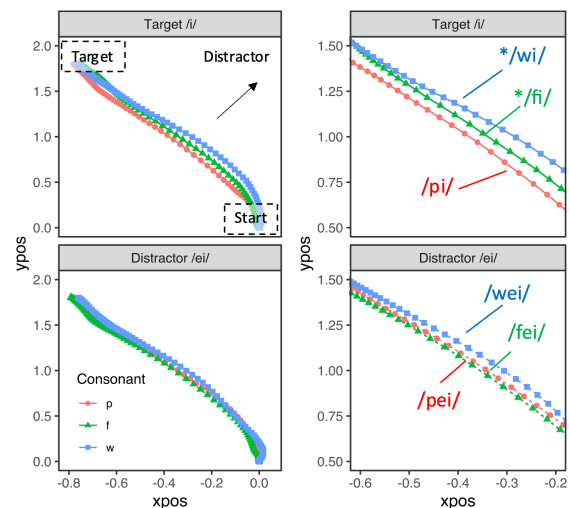


Figure 3. Average mouse trajectories in correct trials. Trajectories are curved towards distractor in illicit contexts. The middle portions are shown on the right.

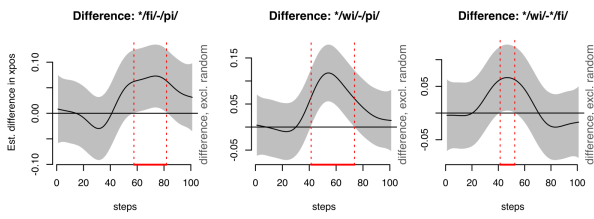


Figure 4. Pairwise contrasts of trajectories showing the differences between average trajectories in the  $x$ -axis between conditions. Positive components indicate movements further towards the distractor /ei/.

## 4. Discussion

The keystroke identification task presented in Section 3.1 provides strong evidence that L1 phonotactic distribution can interfere with L2 vowel perception, as illustrated by the fact that the illicit English category \*/wi/ induces significantly more misidentifications (13.2%) than \*/fi/ (3.7%) and the control condition /pi/ (0.3%), although the difference between \*/fi/ and /pi/ was not significant (Friedman’s test), i.e., \*/wi/ > {\*/fi/, /pi/}. Furthermore, the excellent identification performance in all three distractor conditions indicates that phonological mismatches cause directional perceptual biases. For example, the vowel in /wei/ was rarely misidentified (2.7%), compared to the licit allophone \*/wi/. This type of directional confusion might be understood as the cognitive basis of illusory segment substitutions (\*/wi/ → /wei/, \*/fi/ → /fei/) in cross-language speech perception. Previous research has shown that listeners show directional perceptual biases when perceiving consonants in an L1-illicit context. For instance, Japanese listeners tend to substitute consonants in English CV sequences (e.g., \*/si/ → /ʃi/) due to co-occurrence restrictions [3]. Similar effects were also found in English listeners’ identification of liquid consonants [16]. The results of the present study indicate that vowel substitution can be an alternative strategy to resolve an illicit L2 sequences, apart from epenthesis [4], [5]. The RT data in this experiment shows that RT is longest for \*/wi/, followed by \*/fi/ and /pi/, suggesting that co-occurrence restrictions and contextual allophony in Mandarin phonology [13] have different psycholinguistic properties: both co-occurrence restrictions and contextual allophony can lead to a significant perceptual bias, but to different degrees. Indeed, when compared to the control (licit) condition, the co-occurrence restriction alone has led to an RT increase of 64 ms ( $p < .05$ ), while the contextual allophony has led to an increase of 152 ms ( $p < .0001$ ). The long RTs may indicate an increased amount of evidence accumulation required before reaching the decision threshold due to perceptual biases, i.e., listeners are reluctant to choose vowels which are illicit in a presented context until a substantial amount of evidence accumulates.

The mouse-tracking task presented in Section 3.2 is a methodological extension of the keystroke task, and the %Error patterns observed in this second experiment were highly consistent with those observed in the first. While we found that the RT differences in the mouse-tracking task are less apparent than in the keystroke task, the very similar response speeds between the two experiments in the different conditions justify our further analysis of cursor trajectories along a normalized time course. The mouse cursor trajectories provided rich information about cognitive processes during decision-making [9], especially the level of category competition between the

target and the distractor, which cannot be directly observed in %Error or RT metrics. On the normalized time course, we observed a positive component in the  $x$ -coordinate in the /fi/ and /wi/ conditions, which reflects spurious activation of distractor categories. The competition between \*/fi/ and /fei/ causes a significant trajectory mismatch of 0.073 units in the  $x$ -coordinate (7.3% of the width of the left half-screen), as compared to the vowel competition in the control condition (i.e., /pi/ vs /pei/). Not surprisingly, the competition between \*/wi/ and /wei/ has resulted in a mismatch of 0.117 units, indicating stronger interference. The direct comparison between \*/fi/ and \*/wi/ trajectories also confirmed this difference. According to these results, phonological interference of different strengths will manifest in hand motion during responses and again, contextual allophony results in stronger interference than co-occurrence restrictions alone in L2 vowel perception.

To our knowledge, the present study is the very first study of its kind to examine cross-language speech processing using the mouse-tracking paradigm [7], [9]. As pointed out earlier, %Error and RT alone are limited in revealing the details of the decision-making process during an identification procedure. According to our findings, the keystroke and the mouse-tracking tasks have complementary characteristics. A keystroke task involves relatively simple response processes (i.e., finger pressing), and as such, the RT measure is highly sensitive to the cognitive demands of the task. However, key pressing motions are one-off events, and cannot provide information about continuous interference over time. On the contrary, the mouse-tracking paradigm provides dynamic and readily interpretable information about category competition and decision conflicts with millisecond resolution. It is also possible that the small differences in responses between conditions are due to the difficulty in giving responses, for example, participants must use their hands in coordination with their perception and vision to provide an identification response. At least in the final stage of the response-giving process, some smooth pursuit of the eyes is necessary to guide the hand movement to place the mouse cursor on the selection box. Conversely, pressing a key is a relatively simple motion and does not require the continuous guidance of the eye. Alternatively, the small RT differences may be explained by different decision-making processes employed in the keystroke and mouse-tracking tasks. For instance, participants can easily adjust their choices *en route* during the goal-approaching process, but it is relatively difficult to change their decisions in the keystroke task. It also means that these two paradigms may be used together to allow for triangulation of data while also revealing different aspects of cognitive processing. More broadly, the present study has also shed light on the online cognitive processing underlying phonological repair effects in cross-language speech perception. In particular, perceptual modification of illicit L2 phonemic sequences involves spurious activation of alternative representations that comply with the listeners’ L1 phonology, which leads to a substantial increase in category competition in perception. A further study could assess the implicit discrimination of the vowel contrast in non-native listeners, as well as the learnability of such novel phonotactic regularities in L2 perception.

## 5. Acknowledgements

We want to thank the fifteen Mandarin listeners for their participation. Thanks also go to Alexander Kilpatrick and Jonathon Lum for their help in generating the stimuli.



## 6. References

- [1] C. T. Best, "A direct realist view of cross-language speech perception," in *Speech perception and linguistic experience: Issues in cross-language research*, W. Strange, Ed. Timonium, MD: York Press, 1995, pp. 171–204.
- [2] C. T. Best and M. D. Tyler, "Nonnative and second-language speech perception: Commonalities and complementarities," in *Language experience in second language speech perception*, O.-S. Bohn, Ed. Amsterdam: John Benjamins, 2007, pp. 13–34.
- [3] A. Kilpatrick, R. L. Bundgaard-Nielsen, and B. J. Baker, "Japanese co-occurrence restrictions influence second language perception," *Appl. Psycholinguist.*, vol. 40, no. 2, pp. 585–611, 2019.
- [4] E. Dupoux, E. Parlato, S. Frota, Y. Hirose, and S. Peperkamp, "Where do illusory vowels come from?," *J. Mem. Lang.*, vol. 64, no. 3, pp. 199–210, 2011.
- [5] A. Kilpatrick, S. Kawahara, R. L. Bundgaard-Nielsen, B. J. Baker, and J. Fletcher, "Japanese perceptual epenthesis is modulated by transitional probability," *Lang. Speech*, vol. 64, no. 1, pp. 1–21, 2020.
- [6] J. L. Elman, "Language as a dynamical system," in *Mind as motion: Explorations in the dynamics of cognition*, R. F. Port and T. van Gelder, Eds. Cambridge, MA: The MIT Press, 1995, pp. 194–225.
- [7] M. J. Spivey, M. Grosjean, and G. Knoblich, "Continuous attraction toward phonological competitors," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 29, pp. 10393–10398, 2005.
- [8] A. Weber and A. Cutler, "Lexical competition in non-native spoken-word recognition," *J. Mem. Lang.*, vol. 50, no. 1, pp. 1–25, 2004.
- [9] P. E. Stillman, X. Shen, and M. J. Ferguson, "How mouse-tracking can advance social cognitive theory," *Trends Cogn. Sci.*, vol. 22, no. 6, pp. 531–543, 2018.
- [10] T. A. Farmer, S. A. Cargill, N. C. Hindy, R. Dale, and M. J. Spivey, "Tracking the continuity of language comprehension: Computer mouse trajectories suggest parallel syntactic processing," *Cogn. Sci.*, vol. 31, no. 5, pp. 889–909, 2007.
- [11] E. Hehman, R. M. Stoller, and J. B. Freeman, "Advanced mouse-tracking analytic techniques for enhancing psychological science," *Gr. Process. Intergr. Relations*, vol. 18, no. 3, pp. 384–401, 2015.
- [12] P. Cisek and J. F. Kalaska, "Neural correlates of reaching decisions in dorsal premotor cortex: Specification of multiple direction choices and final selection of action," *Neuron*, vol. 45, no. 3, pp. 801–814, 2005.
- [13] S. Duanmu, *The phonology of Standard Chinese*. Oxford, UK: Oxford University Press, 2007.
- [14] S. Duanmu, "Chinese syllable structure," in *The Blackwell companion to phonology*, vol. 5, M. van Oostendorp, C. J. Ewen, E. Hume, and K. Rice, Eds. Malden, MA: Wiley-Blackwell, 2011, pp. 2151–2177.
- [15] S. Duanmu, "The syllable phonology of Mandarin and Shanghai," *Proc. Fifteenth North Am. Conf. Chinese Linguist.*, pp. 86–102, 2003.
- [16] D. W. Massaro and M. M. Cohen, "Phonological context in speech perception," *Percept. Psychophys.*, vol. 34, no. 4, pp. 338–348, 1983.
- [17] K. Durvasula and J. Kahng, "Illusory vowels in perceptual epenthesis: The role of phonological alternations," *Phonology*, vol. 32, no. 3, pp. 385–416, 2015.
- [18] R. L. Bundgaard-Nielsen, C. T. Best, and M. D. Tyler, "Vocabulary size is associated with second-language vowel perception performance in adult learners," *Stud. Second Lang. Acquis.*, vol. 22, pp. 433–461, 2011.
- [19] G. Stoet, "PsyToolkit: A software package for programming psychological experiments using Linux," *Behav. Res. Methods*, vol. 42, no. 4, pp. 1096–1104, 2010.
- [20] G. Stoet, "PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments," *Teach. Psychol.*, vol. 44, no. 1, pp. 24–31, 2017.
- [21] J. W. Peirce, "PsychoPy: Psychophysics software in Python," *J. Neurosci. Methods*, vol. 162, no. 1–2, pp. 8–13, 2007.
- [22] J. W. Peirce, "Generating stimuli for neuroscience using PsychoPy," *Front. Neuroinform.*, vol. 2, no. 10, 2009.
- [23] M. Wieling, "Analyzing dynamic phonetic data using generalized additive mixed modeling: A tutorial focusing on articulatory differences between L1 and L2 speakers of English," *J. Phon.*, vol. 70, pp. 86–116, 2018.