Interspeech 2025
17-21 August 2025, Rotterdam, The Netherlands

Influence of Proficiency and L2 Experience on Dynamic Spectral Cue Utilization in L2 Vowel Perception and Production

Linda Bakkouche[1], *Brechtje Post*[1]

[1]University of Cambridge, United Kingdom
lb983@cam.ac.uk, bmbp2@cam.ac.uk

Abstract

The acquisition of English vowels as an L2 is complex, yet most studies focus on static measures, with little attention to dynamic spectral cues like Vowel-Inherent Spectral Change (VISC). It remains unclear how language experience and length of residence (LOR) in immersion-rich environments affect perception-production alignment. This study examines Polish learners' perception and production of /e-æ/ (DRESS-TRAP) and /i-ɪ/ (FLEECE-KIT). These contrasts are challenging due to phonetic similarity and category overlap as predicted by L2 models. Advanced learners showed greater perceptual accuracy and more consistent production, especially for /i-ɪ/, while /e-æ/ remained difficult. With higher proficiency, learners exhibited greater formant movement (20–40% of vowel duration), but LOR and language experience were not significant predictors. These findings provide insight into phonetic similarity in theoretical models of L2 vowel acquisition.

Index Terms: Second Language acquisition, Perception-Production Link, Vowel-Inherent Spectral Change (VISC)

1. Introduction

1.1. Background

The acquisition of L2 vowel systems is inherently complex, shaped by the interplay of cross-language phonetic similarities and discrepancies between English and learners' native languages (L1) [1, 2]. For instance, British English vowels that tend to present difficulties for learners include /æ/, /ʌ/, and /ɒ/, which are typologically less common, as well as contrasts like /iː/–/ɪ/ (FLEECE–KIT) and /e/–/æ/ (DRESS–TRAP) [3], which differ in both vowel quality and duration. For learners from L1 backgrounds with limited vowel inventories, such as Polish (/i, ɨ, ɛ, a, ɔ, u/), these vowels and contrasts often lead to assimilation errors, where L2 sounds are mapped onto existing L1 categories [3, 4], e.g., equating English /iː/ and /ɪ/ with Polish /i/, resulting in inaccurate perception and production [5].

Models of L2 speech acquisition, including the (revised) Speech Learning Model (SLM; [1], SLM-r [6]), the Perceptual Assimilation Model (PAM; [2]), and the Second Language Linguistic Perception Model (L2LP; [5]), explain these difficulties by emphasizing the role of cross-language acoustic similarity. The SLM posits that acoustic similarity between L1 and L2 sounds leads to equivalence classification, where learners treat similar sounds as identical, hindering new category formation [1, 6]. PAM focuses on how L2 sounds are perceptually assimilated into existing L1 categories based on their "goodness of fit," predicting which L2 sound contrasts are likely to be confused [2]. While both models emphasize the dominance of L1 categories in early L2 perception, they largely focus on static acoustic representations, limiting their ability to fully explain the acquisition of dynamic vowel cues.

A major challenge in L2 vowel acquisition arises from cross-language phonetic similarity, particularly in relation to the extensive Vowel-Inherent Spectral Change (VISC) in the English vowel system, which refers to formant movement over time [7]. VISC plays a crucial role in distinguishing vowel contrasts such as /e/–/æ/ and /iː/–/ɪ/, where static formant differences are minimal, but dynamic formant trajectories remain essential for accurate identification [4]. Despite its significance, research on VISC in L2 acquisition remains limited, with most studies focusing on foreign language (FL) learning contexts where exposure to naturalistic L2 input is restricted [3].

Moreover, there is limited understanding of the extent to which L2 experience, such as length of residence (LOR) and frequency of use, contributes to the development of sensitivity to dynamic vowel cues. While the L2LP model suggests that prolonged exposure to high-quality L2 input fosters perceptual recalibration, particularly in immersive settings where learners frequently encounter dynamic acoustic cues, empirical research on how these experiential factors shape sensitivity to VISC remains scarce. Understanding these influences is crucial, as perceptual adaptation not only enhances vowel identification but also leads to improvements in production, reinforcing the perception–production link.

This study examines how proficiency, length of residence (LOR), and L2 frequency of use influence Polish learners' perception and production of challenging English vowel contrasts (/e, æ, iː, ɪ/) compared to their Polish counterparts (/ɛ, a, i, ɨ/). By focusing on dynamic acoustic measures, it investigates the relationship between perceptual accuracy and production, providing deeper insight into how immersive exposure, along with LOR and L2 usage shapes vowel acquisition in a second language, proposing that dynamic spectral cues and learner experience are crucial factors that current models have largely overlooked.

2. Methodology

The experiment consisted of two tasks: a perception task testing L2 identification accuracy of English vowel contrasts, and a production task examining the same vowels. Both tasks were conducted in person in a quiet room, starting with a questionnaire, followed by the perception task, and ending with the production task.

Ninety participants participated: 33 intermediate and 32 advanced Polish learners of English, along with 25 native Southern Standard British English (SSBE) speakers as a control group. The intermediate group had a mean age of 31.27 years (SD = 4.55), with a mean age on arrival (AoA) of 22.65 years (SD =

10.21437/Interspeech.2025-954

2.91), and a mean length of residence (LOR) of 8.73 years (SD = 3.93). The advanced group had a mean age of 34.15 years (SD = 5.32), with an AoA of 24.10 years (SD = 3.12) and a mean LOR of 10.50 years (SD = 4.23).

The intermediate group used the second language 74.32% (SD = 16.42) at home, 85.41% (SD = 13.25) in social settings, and 90.81% (SD = 11.40) at work/school, while the advanced group reported 82.50% (SD = 18.54) at home, 90.20% (SD = 14.30) socially, and 95.45% (SD = 12.98) at work/school.

2.1. Stimuli and Experiment Procedure

For the perception task, 288 tokens of monosyllabic minimal pairs (e.g., *sheep* vs. *ship*, *bed* vs. *bad*) were recorded by four native Southern Standard British English (SSBE) speakers (two male, two female). The stimuli were presented through headphones in a randomized order using a two-alternative forced-choice identification task, where participants selected the correct word for each stimulus, thus avoiding order effects.

For the production task, participants produced the target vowels /iː/, /ɪ/, /e/, and /æ/ in English target words matched with the perception task, embedded in the carrier sentence *"I say . . . this time"*. The same participants produced phonologically matched Polish target words with the equivalent Polish vowels /ɛ/, /a/, /i/, and /ɨ/.

2.2. Data Analysis

Data from the perception and production tasks were analyzed separately using R (Version 4.2.1) [8]. Although a broader range of vowel contrasts was included to ensure variability and reduce participant strategy effects, this analysis focuses on the vowel contrasts that are phonetically similar to Polish vowels: /i–ɪ/ (FLEECE–KIT) and /e–æ/ (DRESS–TRAP). Group-level differences and interactions between vowel contrasts and proficiency were explored using two linear mixed-effects models from the `lme4` package [8], with Group (Intermediate, Advanced), Vowel Contrast, and their interaction as fixed effects, and Participant as a random intercept (Bonferroni corrected).

For production data, both static and dynamic aspects of vowel articulation were analyzed. Vowel-Inherent Spectral Change (VISC), a dynamic spectral feature, was selected as the key methodological framework because it is critical for distinguishing vowel identity in English vowels. VISC was measured between the 20% and 80% intervals of each vowel token to capture its dynamic spectral movement [3, 7]. This allowed for the inclusion of both steady-state targets and dynamic cues in the analysis. F1–F2 Euclidean distances (in Bark) served as the primary measure of production accuracy, with smaller distances indicating greater alignment with native vowel targets. Each Euclidean distance was calculated between a learner's F1–F2 values and the mean F1–F2 values of native SSBE speakers for the corresponding vowel and interval. For native SSBE tokens, distances were computed relative to the native group mean, providing a baseline for comparison. Temporal dynamics were modeled by measuring these distances at the second (20–40%), third (40–60%), and fourth (60–80%) intervals, excluding the first and fifth intervals due to acoustic influences from adjacent consonants [3, 7]. Vowel duration was standardized by vowel type to account for temporal differences [9]. Complementary indicators of vowel target positions, such as mean F1-F0 and F3-F2 Bark differences, were also extracted [10].Vowel duration was standardized by vowel type to account for temporal differences [9]. However, since these additional measures did not yield significant effects they are not analyzed further here.

Separate models were fitted for TRAP–/æ/, DRESS–/e/, FLEECE–/i/, and KIT–/ɪ/ to capture contrast-specific variability. These models examined how different vowel contrasts were produced across groups. Euclidean distance measures were analyzed using mixed-effects models in R, with GROUP and GROUP:INTERVAL interactions as fixed effects, and INTERVAL (levels 2, 3, and 4) modeling temporal dynamics. Vowel duration (centered and standardized) was included as a continuous predictor, and random effects accounted for lexical and speaker variation. Significance was tested using *Satterthwaite's* method [11], and model diagnostics confirmed validity.

To explore temporal group-level differences in formant movement, Generalized Additive Models (GAMs) were applied using the `mgcv` package in R [8, 12], with Group (L1 Polish, Intermediate, Advanced, SSBE) as a fixed effect, and smooth terms for interval and duration. Tensor product interactions were used to capture the interaction between duration and interval. By-group smooth terms for INTERVAL allowed for significance testing of group effects [13]. Models accounted for residual autocorrelation [13], and significance was determined by comparing models with and without group-level terms using the `itsadug` package [14].

To analyze the perception–production link, we used linear mixed-effects modeling with the `lme4` package in R [8]. The models included Perception Accuracy, Production Accuracy, Group (Intermediate, Advanced), and Vowel Contrast as fixed effects, with interactions capturing their dynamic relationship across proficiency levels. Participant and Vowel Pairs were modeled as random intercepts to account for individual and vowel-specific variability. Separate models were fitted for each vowel contrast: TRAP–/æ/–DRESS–/e/, FLEECE–/i/–KIT–/ɪ/, to capture contrast-specific variability in the perception–production link. Correlation analyses were conducted within each proficiency group for each vowel to examine the perception–production relationship. Post-hoc comparisons with Bonferroni corrections further identified group-specific differences, offering a detailed understanding of how perceptual gains translated into production accuracy.

3. Results

3.1. Vowel Perception Accuracy

The results indicate that proficiency level plays a crucial role in enhancing perceptual accuracy, particularly for more challenging vowel contrasts, as shown in Figure 1. The mixed-effects analysis further confirmed that proficiency was the primary factor influencing vowel perception. Specifically, for the /e-æ/ (DRESS–TRAP) contrast, advanced learners exhibited significantly higher perceptual accuracy compared to intermediate learners (β = 0.537, SE = 0.147, t = 3.619, p < .001). LOR had a small but significant effect (β = 0.072, SE = 0.031, t = 2.285, p = 0.023), comparable to Language Use (β = 0.064, SE = 0.028, t = 2.247, p = 0.025). Despite these influences, /æ/ remained the most challenging vowel, with intermediate learners averaging 68.3% accuracy, compared to 77.2% for advanced learners. By comparison, the DRESS vowel (/e/) was also difficult for intermediate learners (70.1% accuracy), while advanced learners achieved 81.4% (β = 0.486, SE = 0.142, t = 3.408, p = 0.001). LOR (β = 0.063, SE = 0.032, t = 1.945, p = 0.054) and Language Use (β = 0.058, SE = 0.026, t = 2.276, p = 0.024) had limited but notable effects, suggesting small experience-related improvements. On the other hand, the KIT vowel (/ɪ/) was comparatively less challenging, with intermediate learners reaching 76.8% accuracy and advanced learners nearing native-like levels (87.6%; β = 0.498, SE = 0.129, t = 3.851, p < .001). LOR (β = 0.068, SE = 0.031, t = 2.232, p = 0.027) and Language Use (β = 0.065, SE = 0.029, t = 2.212, p = 0.029) showed minor but consistent contributions. Similarly, the FLEECE vowel (/iː/) highlighted the advantage for advanced learners, who achieved near-native accuracy (83.5%), outperforming intermediate learners (74.2%; β = 0.612, SE = 0.133, t = 4.614, p < .001). LOR (β = 0.077, SE = 0.029, t = 2.691, p = 0.008) and Language Use (β = 0.065, SE = 0.027, t = 2.437, p = 0.016) modestly enhanced perceptual accuracy.

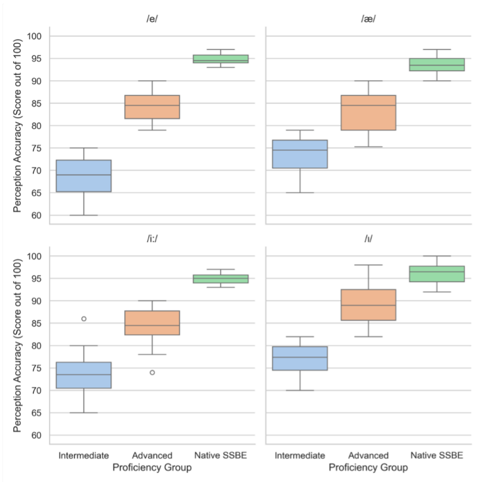

Figure 1: *Perceptual Accuracy (%) Across Vowels and Proficiency Levels.*

3.2. Vowel Production Accuracy

Next, we examined Polish learners' accuracy in producing English vowels by analyzing static formant values and dynamic spectral patterns (Vowel-Inherent Spectral Change, VISC). Baseline data showed that in British English (SSBE), the 20–40% portion of vowel duration, following the CV transition, exhibited the most formant movement, whereas Polish movement occurred later in the vowel. Results suggest that learners recalibrate these time courses to match target-language norms.

Mixed-effects models revealed significant Group × Interval interactions, with advanced learners outperforming intermediate learners. The DRESS–TRAP (/e-æ/) contrast was the most challenging, while the FLEECE–KIT (/iː-ɪ/) contrast posed fewer difficulties. For the TRAP vowel (/æ/), advanced learners exhibited significantly lower Euclidean distances than intermediate learners at Interval 2 (β = 259, p = .004). By Interval 3, both groups had improved, with advanced learners aligned more closely with SSBE norms (β = –617, p < .001) than intermediate learners (β = –654, p < .001). In Interval 4, advanced learners showed more stable production (β = –1916, p < .001) compared to intermediate learners (β = –654, p < .001). For the DRESS vowel (/e/), advanced learners reduced Euclidean distances at Interval 3 (β = 131, p = .012) and further improved at Interval 4 (β = 172, p = .007), while intermediate learners showed no significant change (β = –11, p = .943).

FLEECE–/iː/ posed fewer challenges, with intermediate learners showing a large distance reduction during Interval 3 (β = –1612, p < .001), though advanced learners' productions were more consistent (β = 229, p < .001). KIT–/ɪ/ was the least difficult, with intermediate learners reducing distances significantly in Interval 3 (β = –478, p = .001), while advanced learners maintained closer alignment with SSBE norms and showed minimal changes (β = 91, p = .522).

GAM analyses (Figure 2) indicated greater formant movement during Interval 2 (20–40%) for SSBE speakers in TRAP–/æ/ and DRESS–/e/, while Polish speakers exhibited increased movement in Interval 4 (60–80%). Advanced learners adjusted their vowel trajectories toward native patterns at an earlier point in the vowel (20–40% into the vowel), whereas intermediate learners' adjustments occurred later (60–80%), especially for the non-high vowels. Language Use (β = 297, p = .003) and Length of Residence (LOR; β = 284, p = .002) were significant but secondary to proficiency. TRAP–/æ/ and DRESS–/e/ remained the most difficult contrasts, while FLEECE–/iː/ and KIT–/ɪ/ showed more native-like patterns among advanced learners. Figure 2 focuses only on the TRAP (/æ/) and KIT (/ɪ/) vowels, as these represent opposite ends of the difficulty spectrum for Polish learners—TRAP being the most challenging and KIT the least—and thus illustrate representative contrasts without redundancy, given that DRESS and FLEECE exhibited similar patterns. Individual variation was notable, with some advanced learners achieving near-native trajectories while others showed persistent divergence.

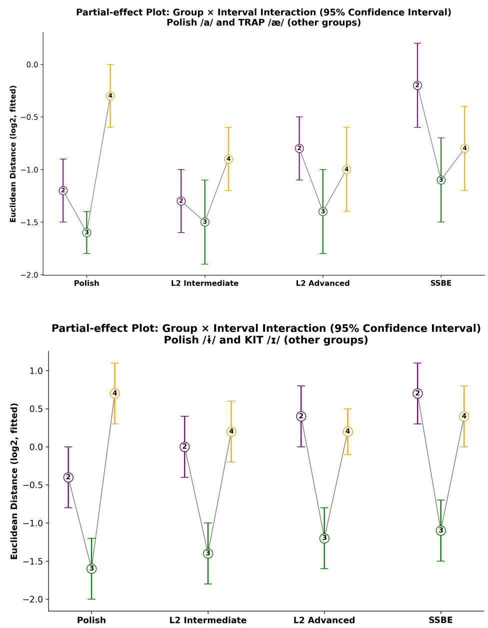

Figure 2: *95% Confidence Interval Partial Effect Plots of GROUP × INTERVAL Interactions. Top: TRAP (/æ/). Bottom: KIT (/ɪ/).*

3.3. Perception-Production Link

To verify whether accurate perception leads to accurate production, linear mixed-effects models were constructed with Perception Accuracy as a key predictor of Production Accuracy, including Group (Intermediate, Advanced) and Vowel Contrast as fixed effects. A significant main effect of Perception Accuracy (β = 0.472, p < .001) indicated that higher perceptual accuracy was associated with more precise production. A significant Perception Accuracy × Group interaction (β = 0.351, p = .002) showed that this relationship was stronger for advanced learners compared to intermediate learners. To account for the role of experience-related factors, interactions involving both Perception Accuracy and Production Accuracy were examined. Specifically, significant interactions between Perception Accuracy × LOR (β = 0.104, p = .018) and Perception Accuracy × Language Use (β = 0.092, p = .024) indicated that longer residence and greater L2 use enhanced the perception–production link. These experience-related factors not only strengthened the perception–production relationship but also influenced production accuracy, suggesting that prolonged and frequent L2 exposure facilitates perceptual recalibration, which in turn supports more accurate articulatory outcomes.

To disentangle how perception accuracy influenced production across vowel contrasts, correlation analyses were run within each group and vowel contrast. For the TRAP vowel (/æ/), advanced learners displayed the strongest correlation (r = 0.712, p < .001), suggesting that their improvements translated into accurate production. In contrast, intermediate learners showed a non-significant link (r = 0.457, p = .067), indicating limited alignment between perception and production for this challenging vowel. For the DRESS vowel (/e/), advanced learners exhibited a significant correlation (r = 0.635, p < .001), while intermediate learners demonstrated a weaker but significant link (r = 0.414, p = .015). For less challenging vowels, both groups showed weaker but significant correlations. For FLEECE–/i/, advanced learners (r = 0.574, p < .001) outscored intermediate learners (r = 0.338, p = .028). Similarly, for KIT–/ɪ/, advanced learners (r = 0.516, p <.001) had a stronger correlation than intermediate learners (r = 0.372, p = .021).

Post-hoc comparisons confirmed that the perception–production link was significantly stronger for advanced learners in the more difficult contrasts: TRAP–/æ/ (t = 3.65, p < .001) and DRESS–/e/ (t = 3.08, p = .003). No significant group differences were observed for the less challenging contrasts, FLEECE–/i/ (t = 1.52, p = .121) and KIT–/ɪ/ (t = 1.21, p = .229). These findings highlight proficiency as key to a stronger perception–production link, with advanced learners showing greater alignment for difficult vowel contrasts.

4. Discussion

This study examined how proficiency, length of residence (LOR), and language use influence Polish learners' perception and production of English vowel contrasts (/e, æ, i, ɪ/) compared to their Polish counterparts (/ɛ, a, i, ɨ/). By focusing on dynamic acoustic measures, it aimed to address gaps in understanding how immersive L2 exposure shapes perception and production, particularly for dynamic contrasts involving VISC. Prior research in foreign language (FL) contexts with limited L2 input has left open questions about how greater exposure in immersive settings fosters perceptual recalibration and production improvement [3]. This study also aimed to determine whether increased L2 experience, through LOR and language use, strengthens the perception–production link and how proficiency contributes to this relationship.

The findings of this study provide valuable insights into the complex relationship between perception, production, and proficiency in second language (L2) vowel acquisition, specifically among Polish learners of English. Consistent with the Speech Learning Model (SLM and SLM-r) [1, 6] and the Perceptual Assimilation Model (PAM) [2], learners struggled with perception of phonetically similar contrasts, such as /e–æ/ (DRESS–TRAP) and /i–ɪ/ (FLEECE–KIT), particularly among intermediate learners. However, not all similar contrasts were equally challenging. Advanced learners achieved near-native accuracy for /i–ɪ/, suggesting that the nature of the contrast matters. Vowel quality contrasts like /e–æ/ involve differences in static formant values and dynamic spectral cues—formant movement over time—that are harder for learners to perceive compared to duration-based contrasts like /i–ɪ/ that rely more on vowel length differences [3].

Production results mirrored perception, with advanced learners consistently outperforming intermediate learners. For TRAP–/æ/, advanced learners showed recalibration of formant trajectories in the earlier vowel portion (20–40%) to align more closely with SSBE norms, while intermediate learners showed greater variability and weaker alignment, particularly in the later portion (60–80%) [7]. These results highlight Vowel-Inherent Spectral Change (VISC) as crucial in L1–L2 vowel similarity and progress tracking [3, 7]. The varying difficulty of phonetically similar vowels suggests that contrasts like /i–ɪ/ may be easier to acquire due to perceptual salience and easier articulatory adjustments.

The present findings demonstrate that proficiency is crucial for establishing a strong perception–production link in L2 vowel learning. Advanced learners showed significantly tighter perception–production correlations, implying that robust L2 phonetic representations develop at higher proficiency levels. By contrast, neither LOR nor language use emerged as major predictors of performance, suggesting that other individual factors (e.g., perceptual or cognitive abilities) may play a more prominent role [15]. This aligns with evidence that immersion and L2 experience alone are insufficient to overcome the challenges of highly similar L1–L2 contrasts [16, 17, 18].

These findings challenge models like SLM-r [6] and PAM [2], which often overlook dynamic aspects of vowel acquisition. The recalibration of formant trajectories in advanced learners underscores the need for L2 speech models to incorporate VISC and time-sensitive parameters. The L2LP model aligns with these results [5], emphasizing that prolonged L2 exposure facilitates perceptual adaptation and production improvements. Moreover, the observed variability among learners underscores the need for more comprehensive L2 speech models that incorporate not only dynamic acoustic cues such as VISC but also individual differences in perception and cognition [15], thereby addressing the limitations of current models in capturing the diversity of learners' adaptation trajectories.

5. Conclusion

This study highlights the pivotal role of proficiency in shaping the perception and production of English vowels by Polish learners. While advanced learners demonstrated significant improvements in both domains, challenging vowel contrasts like /e--æ/ continued to pose difficulties, underscoring the influence of L1 phonetic systems [2, 6]. Notably, not all phonetically similar vowel contrasts were equally difficult; some, like /i:--ɪ/, were acquired to near-native-like levels, suggesting that factors beyond phonetic similarity, such as perceptual salience and articulatory accessibility, influence learning outcomes [15, 17]. These findings highlight the importance of refining theoretical models to account for contrast-specific variability and incorporating dynamic spectral measures to better capture the complexity of L2 vowel acquisition. From a pedagogical perspective, targeted training on difficult contrasts and individualized approaches that address learners' specific challenges are critical for fostering native-like L2 speech [15, 18].

6. References

[1] J. E. Flege. Second language speech learning: Theory, findings, and problems. *in Speech Perception and Linguistic Experience: Issues in Cross-Language Research*, York Press, 1995, pp. 233–277.

[2] C. T. Best and M. D. Tyler. Nonnative and second-language speech perception: Theoretical models and data. *in Second Language Speech Learning*, eds. M. J. Munro and O. Bohn, John Benjamins, 2007, pp. 13–34.

[3] G. Schwartz and J. Dzierla. L1 Polish perception of English vowels: Evidence from spectral dynamics. *Phonetics and Phonology International Journal*, vol. 7, no. 2, 2018, pp. 99–115.

[4] R. A. Fox and E. Jacewicz. Vowel inherent spectral change as a measure of phonetic similarity. *Language and Speech*, vol. 52, no. 2–3, 2009, pp. 273–298.

[5] D. Williams and P. Escudero. Influences of listeners' native and other dialects on cross-language vowel perception. *Frontiers in Psychology*, vol. 5, 2014, article 1053.

[6] J. E. Flege and O. Bohn. The revised Speech Learning Model (SLM-r): Theoretical refinements and research findings. *Journal of Second Language Pronunciation*, vol. 7, no. 3, 2021, pp. 306–356.

[7] G. Schwartz et al. Cross-language vowel comparison: Polish and English contrasts. *Poznań Studies in Contemporary Linguistics*, vol. 52, no. 2, 2016, pp. 285–313.

[8] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing, 2024.

[9] N. Flynn and P. Foulkes. Comparing vowel formant normalization methods. *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, 2011.

[10] A. Syrdal and H. Gopal. A perceptual model of vowel recognition based on the auditory representation of American English vowels. *Journal of the Acoustical Society of America*, vol. 79, no. 4, 1986, pp. 1086–1100.

[11] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, vol. 82, no. 13, 2017, pp. 1–26.

[12] S. N. Wood. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society: Series B*, vol. 73, no. 1, 2011, pp. 3–36.

[13] M. Soskuthy. Generalised Additive Mixed Models for Dynamic Analysis in Linguistics: A Practical Introduction. arXiv preprint arXiv:1703.05339, 2017.

[14] J. van Rij, M. Wieling, R. Baayen, and H. van Rijn. itsadug: Interpreting Time Series and Autocorrelated Data Using GAMMs. R package version 2.3, 2017.

[15] L. Bakkouche and K. Saito. Effects of auditory processing, memory, and experience on early and later stages of second language speech learning. *Second Language Research*, in press, 2025.

[16] K. Saito and K. van Poeteren. Revisiting the perception-production link in L2 speech learning: Evidence from Japanese learners' English /r/. *Applied Psycholinguistics*, vol. 42, no. 3, 2021, pp. 567–591.

[17] G. Granena and M. H. Long. Age of onset, length of residence, language aptitude, and ultimate L2 attainment in three linguistic domains. *Second Language Research*, vol. 29, no. 3, 2013, pp. 311–343.

[18] T. M. Derwing, M. J. Munro, and R. I. Thomson. Immersion and second language pronunciation accuracy. *Studies in Second Language Acquisition*, vol. 30, no. 4, 2008, pp. 533–556.