Universiteit
Leiden
The Netherlands

L1 phonetic reversal and L2 phonetic attrition in Japanese–English bilingual returnee children over the course of five years: an acoustic study

Laméris, T.J.; Kubota, M.

Citation
Laméris, T. J., & Kubota, M. (2026). L1 phonetic reversal and L2 phonetic attrition in Japanese–English bilingual returnee children over the course of five years: an acoustic study. *Second Language Research*, 1-15. doi:10.1177/02676583261461940

Version:	Publisher's Version
License:	Creative Commons CC BY 4.0 license
Downloaded from:	https://hdl.handle.net/1887/4308264

Note: To cite this publication please use the final published version (if applicable).

Check for updates

Research Note

second language research

Second Language Research
1–15
© The Author(s) 2026

Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/02676583261461940
journals.sagepub.com/home/slr

L1 phonetic reversal and L2 phonetic attrition in Japanese–English bilingual returnee children over the course of five years: an acoustic study

Tim Joris Laméris[1] **and Maki Kubota**[2]

Abstract

Only a few studies have acoustically examined how first language (L1) and second language (L2) speech develops once bilingual speakers return to their L1 environment after a prolonged stay in an L2 context, and not much is known about the long-term effects of L1 re-immersion on both languages in phonetic terms. In this study, we analyzed acoustic data from 17 Japanese–English returnee bilingual speakers collected over the course of five years of L1 re-immersion. We ask whether these speakers exhibit changes at the phonetic level in their two languages, and whether these changes pattern with L1 reversal and L2 attrition. An examination of voice onset time, vowel formants, and three rhythmic measures showed that speakers exhibited language-specific values for each of the features. However, apart from some shifts in the L1 Japanese vowel space, we did not observe consistent longitudinal changes in these phonetic features, meaning our acoustic data only partially align with L1 reversal and L2 attrition. Where our data do point in this direction, they complement earlier results from a global foreign-accent rating study with the same speakers, and align with other acoustic studies that looked at L1 reversal but over shorter time spans. Our data tentatively suggest that L1 speech reversal in this returnee population emerges relatively quickly and continues over time, while L2 speech attrition unfolds more gradually.

Keywords

attrition, Japanese, longitudinal, phonetic, returnee, reversal

[1]Leiden University, Leiden, the Netherlands
[2]University of Bergen, Bergen, Norway

Corresponding author:
Tim Joris Laméris, Leiden University, Witte Singel 27A, Leiden, 2311 BG, the Netherlands.
Email: t.j.lameris@hum.leidenuniv.nl

1. Introduction

Studies have shown that the first language (L1) and second language (L2) influence each other bidirectionally in bilingual speakers (Baker and Trofimovich, 2005; Mennen, 2004; Turner, 2023), giving rise to models such as the Attrition and Drift in Access, Production and Perception Theory (ADAPPT: De Leeuw and Chang, 2024), which posit that changes in one language will affect the other language as both 'change and adapt at any time in a person's life' (p. 734). The aim of the present study is to examine these changes and adaptations over time of L1 and L2 speech in the context of returnee bilingualism. In a previous study (Laméris et al., 2025), we identified longitudinal changes in global foreign accent (GFA) of Japanese–English bilingual returnee children. We found that, over the course of five years after return from an L2 English-speaking environment to Japan, the returnees' GFA decreased in L1 Japanese and increased in L2 English. We interpreted these findings based on GFA ratings (a subjective measure) as signs of:

- L1 reversal, a term that we use here to denote linguistic changes that result from increased exposure to and use of the (former) heritage L1 after return to the L1 environment (see Flores and Snape, 2021, 2026); and
- L2 attrition, which here denotes changes to the L2 that result from a decrease in exposure to and use of the L2 (see Flores and Snape, 2026: 4–6).

This study investigates whether such changes also occur at the phonetic level by looking at a set of phonetic features (voice onset time, vowel formants, and rhythmic properties), and, if such changes are present, whether they pattern with L1 reversal and L2 attrition.

1.1. Phonetic drift and attrition in bilingual speech

The process in which an L2 affects L1 speech has been described as 'phonetic drift' or 'phonetic attrition'; see Turner (2023). There is no clear distinction between the two, as they 'characterise outcomes of bilingual speech on a continuum' (De Leeuw and Chang, 2024: 724), but drift is said to concern short-term effects following recent L2 input, whereas attrition typically concerns long-term effects following cumulative L2 input (Chang, 2019b). In the present study, we examine Japanese-L1 children born in Japan who spent a significant portion of their formative developmental years (school age) immersed in an L2 English majority language context (Mean length of residence = 4.27 years, SD = 2.28, Range = 2.00–9.47). These children returned to Japan as early teenagers (Mean age = 10.02 years, SD = 1.72, Range = 7.65–12.99), where they enrolled in Japanese schools. We recorded these children at three time points from the moment of re-immersion in the L1 environment: one week after return to Japan (Time 1); one year after (Time 2), and five years after (Time 3). Given these medium-term to long-term periods of L2 immersion and L1 re-immersion, we will henceforth use the term 'attrition' instead of 'drift' in the context of these speakers.

The ways in which phonetic L1 attrition can take place are described in two scenarios by the revised Speech Learning Model, the *SLM-r* (Flege and Bohn, 2021). The first scenario, phonetic category 'assimilation', involves an approximation of L1 and L2 sounds

within a common phonetic space, in the absence of a newly formed L2 category. This may cause an L1 sound to shift toward (assimilate to) an L2 sound (Flege and Bohn, 2021: 42). For example, an increase in voice onset time (VOT) duration for voiceless stops in L1 Japanese following exposure to L2 English could suggest L1-to-L2 assimilation, given that English VOTs are typically longer than Japanese VOTs (Riney et al., 2007). In the second scenario, phonetic category 'dissimilation', newly formed L2 categories shift away from L1 categories to enhance cross-linguistic contrast. An example here would be if VOT values shorten for Japanese and/or lengthen for English. A third scenario, proposed by Turner (2023), is 'tandem drift', in which acoustic properties of both L1 and L2 sounds shift in the same direction. In the example of VOT, tandem drift could be indicated by an increase or decrease in both Japanese and English VOT durations.

The abovementioned attritional processes have been documented in various phonetic production studies in the context of L2 exposure and/or L2 immersion, for instance, during residence abroad or immersive classroom experience (Chang, 2019b; de Leeuw et al., 2012; Saud Alharbi et al., 2023; Stoehr et al., 2017, and others). However, as pointed out by the ADAPPT model (De Leeuw and Chang, 2024: 725), attrition does not mean loss, and it is a process that can be reversed. Yet, only a handful of production studies have looked at phonetic reversal in the context of L1 re-exposure and re-immersion, for instance, after return to the L1 environment (Kartushina and Martin, 2019; Sancier and Fowler, 1997; Tobin et al., 2017; Turner, 2023). While these studies suggest that L1 production (at least partially) 'resets' to L1 norms over the course of re-immersion, they all investigated bilinguals who acquired their L2 (or L3 in the case of Kartushina and Martin, 2019) in adulthood, and examined re-immersion periods that were shorter than one year. In the present study, we investigate re-immersion effects in early-onset bilingual children over the course of five years of L1 re-immersion. In doing so, we hope to shed more light on the long-term nature of L1 reversal in childhood/early teenage years. We present the following research question:

- Do Japanese–English bilingual returnee children exhibit changes in their speech in phonetic terms over the course of L1 re-immersion, and do these changes align with L1 reversal and L2 attrition, as previously found in their global foreign accent?

1.2. Phonetic properties of English and Japanese

In order to investigate phonetic L1 reversal and L2 attrition, we obtained measures of voice onset time (VOT) in syllable-onset prevocalic voiceless stops; vowel height and frontedness; and a set of rhythmic values for English and Japanese. Here, we briefly summarize the language-specific characteristics of these features and describe how changes in them over time could point to L1 reversal and L2 attrition. We assume here that L1 reversal is marked by a shift to Japanese-typical values, and L2 attrition by a shift away from English-typical values.

1.2.1. VOT. In syllable-onset, prevocalic environments, English voiceless stops /p/ /t/ /k/ are aspirated and have long VOTs, typically ranging between 60–100 milliseconds (ms),

whereas the duration of Japanese voiceless stop is intermediate, ranging between 25 to 60 milliseconds (Cho and Ladefoged, 1999; Holliday, 2019; Riney et al., 2007). Riney and Takagi (1999) investigated changes in English-L2 VOT values by 11 adult Japanese-L1 speakers who were exposed to English over a three-year period at a Japanese university. No changes in their VOT were observed over time, which the authors attribute to the relative overlap between intermediate-lag VOTs in Japanese and long-lag VOTs in English. However, higher VOT values in English were associated with lower English foreign-accent scores, suggesting that some individuals had acquired more English-like VOT values. By contrast, a similar study by Hanzawa (2018) with 50 Japanese-L1 university students did find increases in English VOT (suggesting L2 acquisition) over a one-year period of English instruction at a university in Japan. Although we are not aware of any other longitudinal studies that investigated VOT in Japanese–English bilinguals, we expect that, based on other studies that looked at L1 attrition of VOT (Mayr et al., 2012; Saud Alharbi et al., 2023; Stoehr et al., 2017), a decrease in Japanese VOT would signal L1 reversal, and a decrease in English VOT would signal L2 attrition over the course of L1 re-immersion.

1.2.2. Vowels. The Japanese vowel space consists of five distinct qualities: /a e i o ɯ/, which each have a phonemically long equivalent. General American English (the variety that our returnees were exposed to) has a relatively 'dense' system consisting of 10 to 11 vowels (Nishi et al., 2008: 577). Previous studies have suggested various assimilatory patterns in Japanese–English bilingual vowel production that could signal L1 attrition and L2 acquisition; for a recent overview, see Yazawa et al., 2023: 5–8. For instance, a longitudinal study by Oh et al. (2011) investigated production of English vowels by 16 Japanese children (mean age = 9.9 years) living in the US, who were recorded at half a year (time 1), and one-and-a-half year of L2 immersion (time 2). Over this period, the Japanese children exhibited raising (a lowering of the first formant, F1) of English /i/ and /ɑ/ and backing (a lowering of the second formant, F2) of English /ɛ/. The authors interpreted this as L2 acquisition, as the children's English vowels differed from an age-matched English-L1 control group at time 1, but no longer at time 2, suggesting the children had attained English-typical values. Oh et al. (2011) also compared Japanese–English vowel pairs, (/a/–/ɑ/, /i/–/i/, /u/–/u/), and found that cross-linguistic differences that were present at time 1 were still maintained after a year, except for /i/–/i/, where the raising of English /i/ and fronting of Japanese /i/ appeared to have caused a cross-linguistic merger (i.e., assimilation). The authors suggest that this bidirectional movement between L1 Japanese and L2 English vowel qualities highlights the malleability of bilingual children's speech systems under L2 immersion (Oh et al., 2011: 166). The question that is central to the current article is whether such malleability is also observed under L1 re-immersion. In addition to VOT, we therefore also examined longitudinal changes in Japanese vowels /a e i o ɯ/, together with five English vowels /æ ɛ i ɑ ʊ/. These English vowels were selected based on their acoustic and perceptual similarity to their Japanese 'counterparts' (Nishi et al., 2008), and to allow for a comparison with the studies by Yazawa et al. (2023, 2024), who investigated changes in a similar vowel set. In those studies, vowel formant values by 55 adult Japanese–English sequential bilinguals were compared to their English and Japanese nativelikeness scores. They found, for instance,

that a higher English /i/ was associated with higher English nativelikeness, suggesting that a shift toward a higher /i/ indexes L2 acquisition, similar to the assimilatory pattern found in the Japanese–English children reported in Oh et al. (2011). However, when looking at shifts in the Japanese vowel space, they note that only 'some of the results can be straightforwardly explained by L1–L2 segmental assimilation or dissimilation' and that the supposed assimilation-dissimilation patterns are 'inconsistent at best' (Yazawa et al., 2024: 11). Given these findings from the Japanese–English L2 immersion context, we refrain from making specific predictions about how L1 re-immersion and L2 attrition might manifest in longitudinal changes to vowel properties, aside from the possibility that vowel categories that may have assimilated during L2 immersion (Oh et al., 2011) could dissimilate again if speakers revert to L1 norms.

1.2.3. Rhythmic properties. Japanese and English are rhythmically distinct languages, as Japanese is often described as a mora-timed language with a maximum syllable structure of CV(C) (Hasegawa, 2018: 135–153), and English is a stress-timed language with more complex syllable structures (Tajima and Port, 2004). Previous studies have shown that transfer of language-specific rhythmic properties can occur both unidirectionally (L1-to-L2) (Kawase et al., 2025; Li and Post, 2014), as well as bidirectionally in multilingual speakers (Brown-Bousfield and Chang, 2023).[1] For instance, Kawase et al. (2025) compared rhythm metrics in 10 adult English-L1 and 10 Japanese-L1 speakers, and found that the Japanese speech had reduced durational variability in both consonant and vowel intervals compared to English. They also investigated the Japanese speakers' L2 English, and here too found reduced variability in vowel and consonant intervals, suggesting that L1 Japanese rhythm transferred to L2 English. For the present study, we compared English and Japanese rhythm by obtaining three metrics: VarcoV (Dellwo, 2006), nPVI-V (Grabe and Low, 2002) and %V (Ramus et al., 1999).[2] Based on Kawase et al. (2025), we assume that English has higher variability in vowel duration than Japanese (i.e., English > Japanese for VarcoV and nPVI-V) and that the proportion of vowels is lower in English than in Japanese (i.e., English < Japanese for %V). L1 reversal could therefore be signalled by a decrease in VarcoV and nPVI and an increase in %V in Japanese, and L2 attrition by an increase in VarcoV and nPVI and a decrease in %V in English.

2. Methodology

2.1. Speech materials

Speech materials were the same as in Laméris et al. (2025). An overview of the speakers' demographic information and the speech material is available in supplemental material 2.

2.2. Acoustic measurements

The 102 (17 speakers × 3 times × 2 languages) 10-second speech fragments, including the 34 fragments from age-matched monolingual speakers from Laméris et al. (2025), were annotated in *Praat* (Boersma and Weenink, 2019). The Montreal Forced Aligner (McAuliffe et al., 2017) was first used to automatically annotate the fragments, after

which a Japanese–English bilingual phonetician manually checked the generated TextGrids and further labelled syllable-onset or mora-onset stops for their closure and VOT duration, following the protocol described in Idemaru et al. (2019: 341). F1 and F2 measures were obtained using a *Praat* script and measured at the vowel midpoint. Rhythmic properties were obtained using the *rhythm.metrics* package (Zhang, 2022).

The speech fragments, TextGrids, and scripts and code to generate the acoustic data (including details on outlier exclusion) are available in supplemental material 1.

2.3. Statistical analysis

We ran Bayesian logistic regression models in the *brms* package (Bürkner, 2018) in *R 4.4.1* (R Core Team, 2024) to investigate longitudinal changes for each phonetic feature in the returnees' speech. The model structure was motivated by our research questions[3] and contained fixed effects for *Time* (1, 2, 3, sum contrast-coded); *Language* (English, Japanese, sum contrast-coded), and a *Time*Language* interaction. The random-effects structure consisted of a by-subject random slope for the *Time*Language* interaction. The model for VOT contained an additional random intercept for label (/p/, /t/, /k/). Since we examined a set of language-specific vowels per language, we ran models per vowel in each language. The structure for these models, therefore, only contained a fixed effect for *Time* and a by-subject random slope for *Time*.[4] All models were fitted with a shifted lognormal distribution in order to run analyses on the raw, non-transformed data (Nicenboim et al., 2018: 1078–1079). We used weakly informative priors with prior specification set as (0,1) for 'Intercept', (0,3) for 'b', (0,0.1) for SD priors and (0,1) for sigma priors, and ran four sampling chains with every 3,000 iterations and 1,500 warm-up iterations. Model diagnostics were carried out by observing Rhat values (ensuring these were close to 1) and by inspecting posterior draws using the *brms* pp_ check function. To investigate interactions, planned comparisons were carried out with the *emmeans* package (Lenth, 2020). We report findings for which the 95% Credible Intervals (CrIs) of the effect estimates and the 95% Highest Posterior Density for planned comparison estimates did not contain zero, and take such findings to be 'suggestions' of an effect (Nicenboim et al., 2018).

We also exploratorily investigated if variation in phonetic features was related to variation in Global Foreign Accent as reported in our previous study (Laméris et al., 2025). Here, we followed the statistical approach described in Idemaru et al. (2021) for VOT and rhythmic properties, and Yazawa et al. (2024) for vowels (for details, see supplemental material 1). For brevity, we do not report these findings in the Results section, but we briefly refer to them in the discussion.

3. Results

3.1. VOT

Figure 1 shows VOT (in ms) per language and time of return. Values from age-matched monolingual speakers are shown on the left for reference. A visual inspection suggests that returnees had longer VOT values for voiceless stops in English than in Japanese, but

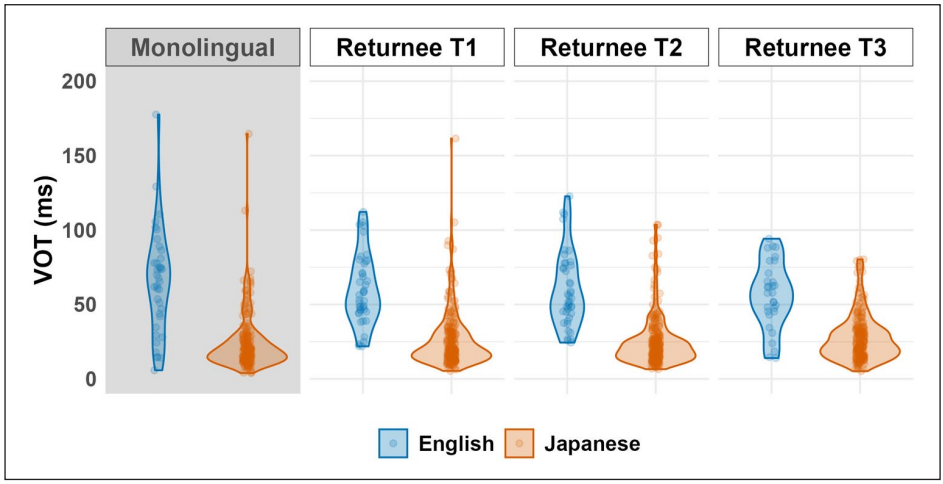

Figure 1. Violin plots displaying VOTs per language and time of return.
Note. Monolingual data are shown on the left for reference.

no clear changes over time can be observed. Indeed, the model for VOT (749 observations) only suggested an effect of *Language*, $b = 0.47$ (0.40, 0.54), and pairwise comparisons suggested that overall, English VOTs were longer than Japanese VOTs, $b = 0.94$ (0.79, 1.08).

3.2. Vowel formants: F1 and F2

Figure 2 shows normalized vowel spaces per language and time of return, with monolingual vowel spaces shown on the left for reference. The models examining changes in F1 and F2 per vowel suggested several shifts over time in vowel height and frontedness. These shifts are summarized in Table 1, which shows the estimates of the pairwise comparisons for *Time* for the models that suggested an effect of *Time*. For instance, in the first row, the comparison for F1 changes in /ɛ/ between Time 1 and Time 2 suggests no change because the 95% HPD range includes zero. However, the comparison between Time 2 and Time 3 suggests a change in F1, as zero is not included in the 95% HPD. The negative estimate sign suggests a decrease in F1, i.e., vowel raising. Based on these results, it appears that only two English vowels (/ɛ/ and /ʊ/) shifted (raised), but that all Japanese vowels exhibited some shifts, consisting of raising, fronting, or a combination thereof.

3.3. Rhythmic measures: VarcoV, nPVI-V, and %V

Figure 3 shows VarcoV, nPVI-V, and %V values per language and time of return, with monolingual values on the left for reference. A visual inspection suggests differences in the values between languages, but no clear changes in these values over time. Indeed, none of the models (each: 102 observations) suggested an effect of *Time* or a *Language:Time* interaction: The VarcoV model suggested a *Language* effect, $b = 0.07$

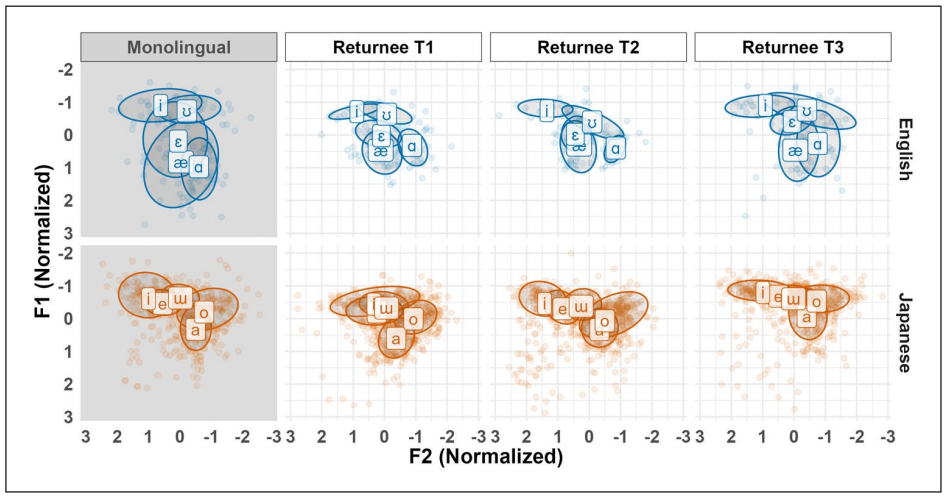

Figure 2. Scatterplots displaying Lobanov-normalized F1 and F2 values per language and time of return.

Table 1. Planned comparisons for Time for F1 and F2 values per vowel.

F1/F2 and vowel	Number of tokens	Estimate (95% HPD): T1 to T2	Estimate (95% HPD): T2 to T3	Inference
English:				
F1.ɛ	77	−0.03 (−0.19, 0.10)	−0.24 (−0.42, −0.09)	Raising of /ɛ/ from T2–T3
F1.ʊ	32	0.25 (−0.18, 0.74)	−0.54 (−1.08, −0.10)	Raising of /ʊ/ from T2–T3
Japanese:				
F1.a	729	−0.10 (−0.17, −0.03)	−0.22 (−0.30, −0.12)	Consistent raising of /a/
F1.e	288	−0.07 (−0.18, 0.02)	−0.29 (−0.40, −0.19)	Raising of /e/ from T2–T3
F2.e	288	0.19 (0.11, 0.27)	−0.09 (−0.16, −0.02)	Unclear*, overall fronting
F1.i	373	−0.05 (−0.18, 0.06)	−0.27 (−0.42, −0.12)	Raising of /i/ from T2–T3
F2.i	373	0.28 (0.19, 0.37)	−0.10 (−0.22, 0.00)	Fronting of /i/ from T1–T2
F1.o	644	−0.01 (−0.13, 0.10)	−0.37 (−0.42, −0.23)	Raising of /o/ from T2–T3
F2.o	644	0.12 (0.01, 0.23)	−0.06 (−0.20, 0.07)	Fronting of /o/ from T1–T2
F1.u	190	−0.06 (−0.17, 0.05)	−0.14 (−0.25, −0.01)	Raising of /u/ from T2–T3
F2.u	190	0.15 (0.03, 0.28)	−0.07 (−0.14, 0.00)	Fronting of /u/ from T1–T2

Notes. Only comparisons from models for which there was a suggestion for a main effect of Time are listed here. * The comparison between T3 and T1 suggested that F2 for /e/ increased, $b = 0.10$ (0.00, 0.19).

(0.00, 0.16), with English values higher than Japanese $b = 0.13$ (0.00, 0.30). Similarly, the nPVI-V model suggested a *Language* effect, $b = 0.57$ (0.35, 0.79), with English values higher than Japanese, $b = 1.13$ (0.69, 1.58). The model for %V also suggested a *Language* effect, $b = -0.19$ (−0.25, −0.16) with English values lower than Japanese, $b = -0.38$ (−0.48, −0.31).

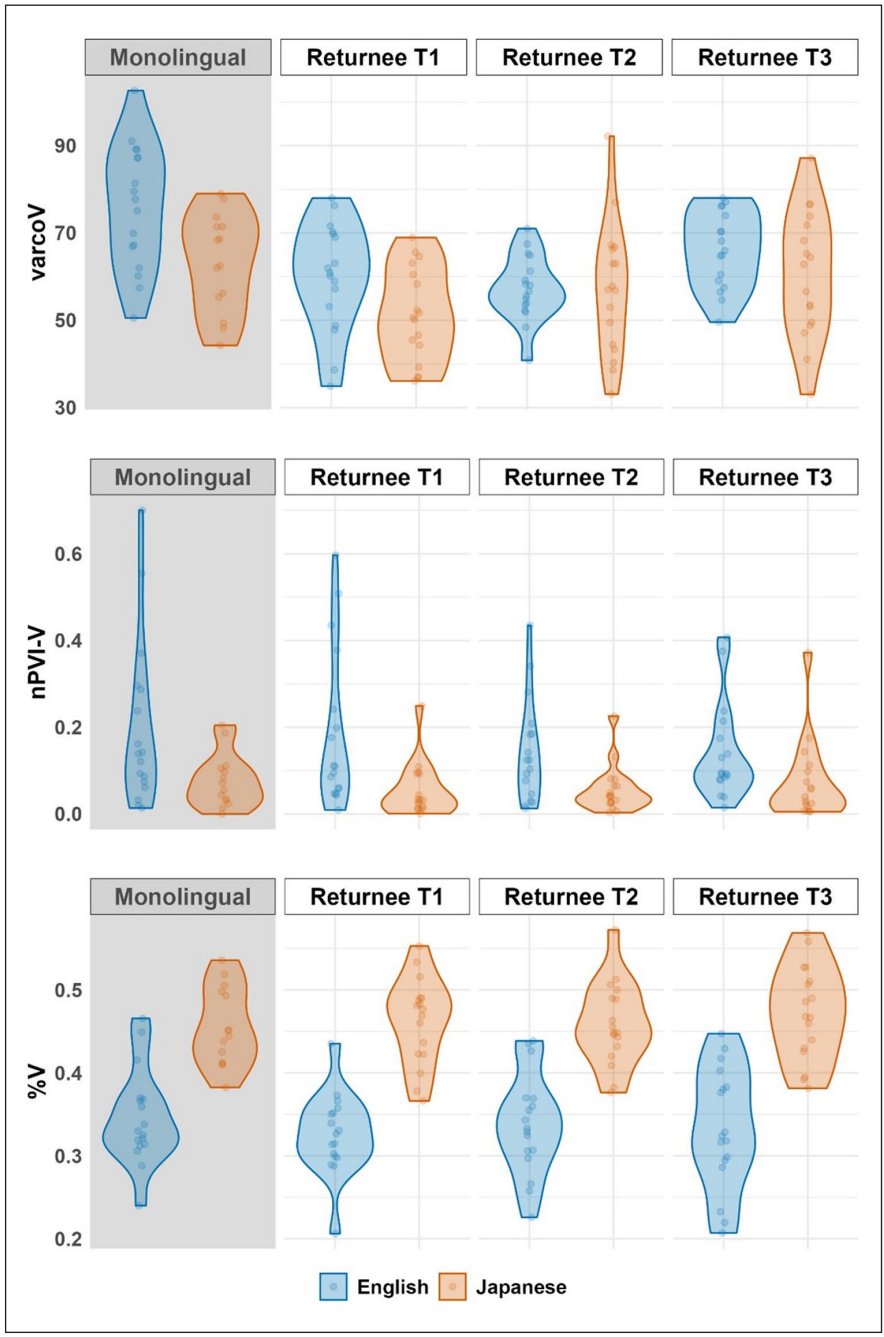

Figure 3. Violin plots displaying VarcoV, nPVI-V, and %V per language and time of return.
Note. Monolingual data are shown on the left for reference.

4. Discussion

Based on a previous study in which we found a decline in Global Foreign Accent (GFA) in L1 Japanese and an increase in GFA in L2 English in Japanese–English bilingual returnee children, we asked in the present study whether similar signs of L1 reversal and L2 attrition could be captured in phonetic terms. Acoustic analyses on the same speech samples as in our accent-rating study yielded limited evidence for phonetic changes in the returnees' English and Japanese speech over the course of five years. Although we found language-typical values for each feature, e.g., long VOTs in English vs. intermediate VOTs in Japanese (Riney et al., 2007), as well as in terms of VarcoV, nPVI-V, and %V (Kawase et al., 2025), the data suggest that these values did not change over time, in either language.

We did find changes over time for vowel height and frontedness. For L1 Japanese, all five vowels /a e i o ɯ/ exhibited raising and/or fronting. For instance, /a/ exhibited consistent raising from Time 1 to Time 2 and from Time 2 to Time 3. In our exploratory analysis that investigated the link between vowel formants and global foreign accent (supplemental material), as well as in the study by Yazawa et al. (2024), a higher Japanese /a/ was associated with a weaker L1 foreign accent (i.e., more native-like). Based on this association, a higher /a/ over the course of five years of L1 re-immersion could signal phonetic L1 reversal. By the same account, raising of /e/ and /o/ was associated with a weaker foreign accent and could similarly signal L1 reversal.[5]

As for English L2, we observed raising of /ɛ/, which has been associated with less-nativelike pronunciations in L2 English (Yazawa et al., 2023) and could therefore signal L2 attrition, but we are cautious to derive strong conclusions given that we only had 77 observations for /ɛ/. A similar concern holds for /ʊ/, for which we only had 33 observations.

The apparent unidirectional shift (upwards and forwards) of Japanese /a e i o ɯ/ could signal a dissimilatory shift away from its English counterparts. For instance, a visual inspection of Japanese /a/ vis-à-vis English /æ/ in Figure 4 suggests that /a/ is relatively close in terms of vowel height to /æ/ at Times 1–2. But by Time 3, /a/ is distinctly higher,[6] and resembles the contrast with /æ/ in the monolingual English–Japanese vowel spaces. Similar observations can be made for the raising of Japanese /e/ and /o/ relative to their English counterparts /ɛ/ and /ɑ/. Thus, perhaps some L1 dissimilatory shift, in line with L1 reversal, did take place for these vowels.

Despite the abovementioned changes in the vowel space, we take these acoustic data as limited evidence for L1 reversal and L2 attrition, given that we did not observe changes in VOT, as reported in other cross-linguistic studies (Chang, 2019a; Hanzawa, 2018, but see Riney and Takagi, 1999), nor did we observe longitudinal changes in rhythmic features. One limitation of our speech material is that it consisted of spontaneous speech (picture narrations), from which only a limited number of usable feature tokens could be extracted per speaker per time. This may have masked longitudinal changes that could have emerged if we had analyzed the exact same target speech material (e.g., target word lists) at different time points (Chang, 2019a; Turner, 2023). In addition, it is possible that the previously reported changes in perceived global foreign accent (Laméris et al., 2025) may reflect changes in phonetic features beyond those examined here, such

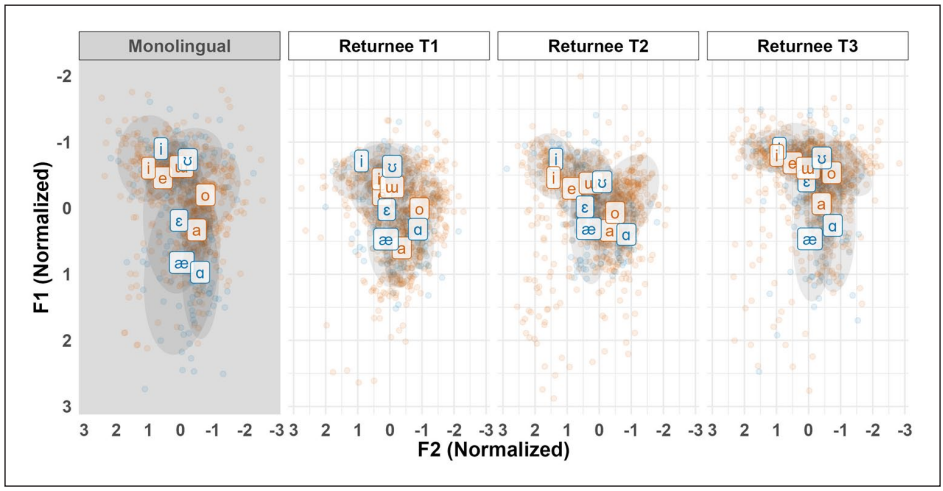

Figure 4. Scatterplots displaying Lobanov-normalized F1 and F2 values per time of return, overlapped for English (blue) and Japanese (red).

as second and third formant values for English liquids (Nagamine, 2024), or lexical pitch-accent targets for Japanese (Idemaru et al., 2019).

In sum, based on an analysis of the acoustic features that we examined in the present study, it appears that some changes in the L1 vowel space point to L1 reversal in Japanese–English returnee children. L2 attrition appears to be less clearly identifiable in phonetic terms. To some extent, this conclusion aligns with Turner (2023), who found that English–French adult sequential bilinguals exhibited partial L1 reversal but no clear L2 attrition (i.e., L2 maintenance) in their vowel space over a relatively short period of L1 re-immersion (10 months). Taken together with the results from our earlier accent-rating study (Laméris et al., 2025), the present study's acoustic data suggest that, in the long term (five years of re-immersion), this pattern of L1 reversal continues, whereas L2 attrition progresses more gradually.

Acknowledgments

The authors would like to acknowledge Satsuki Kurokawa for his help with the annotation of the audio fragments and with the data transcription.

ORCID iDs

Tim Joris Laméris https://orcid.org/0000-0002-1365-3022
Maki Kubota https://orcid.org/0000-0003-0580-1045

Ethical considerations

The University of Edinburgh Linguistics and English Language Ethics Committee (protocol number 11-1516/5) approved this study.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Maki Kubota received an AcqVA Aurora Center grant for this research (UiT The Arctic University of Norway).

Declaration of conflicting interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data availability statement

Data for this research is publicly available at https://osf.io/fb7mj/?view_only=750db21af7944d56af19de81df627313

Supplemental material

Supplemental material for this article is available at https://osf.io/fb7mj/?view_only=750db21af7944d56af19de81df627313

Notes

1. Specifically, in Brown-Bousfield and Chang (2023), an influence from L3 (Spanish) onto L1 (English) was found in terms of VarcoV.
2. These rhythm metrics are defined as follows (Kawase et al., 2025: 2):

   VarcoV: 'The standard deviation of the vowel durations divided by mean vowel duration, multiplied by 100'
   %V: 'Sum of vocalic intervals divided by the total duration of sentences'
   NPVI-V: 'Normalized pairwise variability index for vocalic intervals: Mean of the differences between successive vocalic intervals divided by sum, multiplied for 100'
3. As pointed out by a reviewer, the returnees exhibited considerable heterogeneity in terms of linguistic-experiential factors (country of residence, age of L2 acquisition, exposure to English, and length of residence abroad). We ran exploratory models for VOT that included these variables (see supplemental material: code_returnee_acoustic). These models suggested a role for individual factors, e.g., individuals with more L2 English exposure had longer VOTs (more English-typical) than individuals with less L2 exposure. However, they did not reveal different results regarding the differences in VOT between time and language than the models without the individual factor variables. For brevity and ease of discussion, we therefore report the models that exclude these individual factors, but acknowledge that participant heterogeneity can play a role in phonetic changes in the L1 and L2.
4. Model structure was as follows:

VOT:	duration ~ time*language + (time*language | speaker) + (1 | label)
F1/F2 (per vowel per language):	formant ~ time + (time | speaker)
Rhythmic values:	value ~ time*language + (time*language | speaker)
5. However, in contrast to our findings, in Yazawa et al. (2024) raised /e/, /i/ and /ɯ/ were associated with stronger L1 foreign accents (see supplemental material: code_slr_returnee_acoustic). The rising of these vowels would therefore signal L1 attrition and not L1 reversal. Our exploratory analysis did not suggest that changes in /ɯ/ influenced accentedness scores.

6. To explore if this was indeed the case, we ran an extra Bayesian model that included a *Time*Vowel* interaction on the combined F1 data for /a/ and /æ/ (836 observations). This model suggested that whereas there was no difference in F1 between /a/and /æ/ for Times 1 and 2, there was in fact for Time 3, $b = -0.24$ (−0.34, −0.14).

References

Baker W and Trofimovich P (2005) Interaction of native- and second-language vowel system(s) in early and late bilinguals. *Language and Speech* 48(1): 1–27. https://doi.org/10.1177/00238309050480010101

Boersma P and Weenink D (2019). *Praat: Doing Phonetics by Computer* [computer program]. Available at: https://www.praat.org (accessed June 2026).

Brown-Bousfield MM and Chang CB (2023) Regressive cross-linguistic influence in multilingual speech rhythm: The role of language similarity. In: Brown MM, Flynn S and Fernández-Berkes É (eds) *Studies in Bilingualism*, vol. 65. John Benjamins, pp.49–71. https://doi.org/10.1075/sibil.65.03bro

Bürkner P-C (2018) Advanced Bayesian multilevel modeling with the R package brms. *The R Journal* 10(1): 395–411. https://doi.org/10.32614/RJ-2018-017

Chang CB (2019a) Language change and linguistic inquiry in a world of multicompetence: Sustained phonetic drift and its implications for behavioral linguistic research. *Journal of Phonetics* 74: 96–113. https://doi.org/10.1016/j.wocn.2019.03.001

Chang CB (2019b) Phonetic drift. In: Chang CB (ed.) *The Oxford Handbook of Language Attrition*. Oxford University Press, pp.190–203. https://doi.org/10.1093/oxfordhb/9780198793595.013.16

Cho T and Ladefoged P (1999) Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics* 27(2): 207–229. https://doi.org/10.1006/jpho.1999.0094

De Leeuw E and Chang CB (2024) Phonetic and phonological L1 attrition and drift in bilingual speech. In: Amengual M (ed.) *The Cambridge Handbook of Bilingual Phonetics and Phonology*. Cambridge University Press, pp.721–745.

de Leeuw E, Mennen I and Scobbie JM (2012) Singing a different tune in your native language: First language attrition of prosody. *International Journal of Bilingualism* 16(1): 101–116. https://doi.org/10.1177/1367006911405576

Dellwo V (2006) Rhythm and speech rate: A variation coefficient for deltaC. In: Karnowski P and Szigeti I (eds) *Language and Language-Processing*. Peter Lang, pp.231–241.

Flege JE and Bohn O-S (2021) The revised speech learning model (SLM-r). In: Wayland R (ed.) *Second Language Speech Learning*. Cambridge University Press, pp.3–83. https://doi.org/10.1017/9781108886901.002

Flores C and Snape N (2021) Language attrition and heritage language reversal in returnees. In: Montrul S and Polinsky M (eds) *The Cambridge Handbook of Heritage Languages and Linguistics*. Cambridge University Press, pp.351–372. https://doi.org/10.1017/9781108766340.016

Flores C and Snape N (2026) What returnee bilinguals may teach us about language attrition, language stabilization, and individual variation. *Linguistic Approaches to Bilingualism* 16(1): 1–25. https://doi.org/10.1075/lab.25021.flo

Grabe E and Low EL (2002) Durational variability in speech and the Rhythm Class Hypothesis. *Laboratory Phonology* 7: 515–546. https://doi.org/10.1515/9783110197105.2.515

Hanzawa K (2018) The development of voice onset time (VOT) in a content-based instruction university program by Japanese learners of English: A longitudinal study. *The Canadian Modern Language Review* 74(4): 502–522. https://doi.org/10.3138/cmlr.2018-0196

Hasegawa Y (ed.) (2018) *The Cambridge Handbook of Japanese Linguistics*. Cambridge University Press. https://doi.org/10.1017/9781316884461

Holliday JJ (2019) The perception and production of word-initial Korean stops by native speakers of Japanese. *Language and Speech* 62(3): 494–508. https://doi.org/10.1177/0023830918785649

Idemaru K, Kato M and Tsukada K (2021) Foreign accent in L2 Japanese. In: Wayland R (ed.) *Second Language Speech Learning*. Cambridge University Press, pp.377–396. https://doi.org/10.1017/9781108886901.016

Idemaru K, Wei P and Gubbins L (2019) Acoustic sources of accent in second language Japanese speech. *Language and Speech* 62(2): 333–357. https://doi.org/10.1177/0023830918773118

Kartushina N and Martin CD (2019) Third-language learning affects bilinguals' production in both their native languages: A longitudinal study of dynamic changes in L1, L2 and L3 vowel production. *Journal of Phonetics* 77: Article 100920. https://doi.org/10.1016/j.wocn.2019.100920

Kawase S, Davis C and Kim J (2025) Impact of Japanese L1 Rhythm on English L2 speech. *Language and Speech* 68(1): 118–140. https://doi.org/10.1177/00238309241247210

Laméris TJ, Kubota M, Kupisch T, et al. (2025) Language change in Japanese–English bilingual returnee children over the course of five years: Evidence from accent-rating. *Second Language Research* 41(1): 191–216. https://doi.org/10.1177/02676583241230854

Lenth R (2020) Emmeans: Estimated Marginal Means, aka Least-Squares Means. Available at: https://cran.r-project.org/package=emmeans (accessed June 2026).

Lenth RV, Bolker B, Buerkner P, et al. (2020) *Emmeans: Estimated Marginal Means, aka Least-Squares Means: R package version 1.4.7*. Available at: https://cran.r-project.org/package=emmeans (accessed June 2026).

Li A and Post B (2014) L2 acquisition of prosodic properties of rhythm. *Studies in Second Language Acquisition* 36(2): 223–255. https://doi.org/10.1017/S0272263113000752

Mayr R, Price S and Mennen I (2012) First language attrition in the speech of Dutch–English bilinguals: The case of monozygotic twin sisters. *Bilingualism: Language and Cognition* 15(4): 687–700. https://doi.org/10.1017/s136672891100071x

McAuliffe M, Socolof M, Mihuc S, et al. (2017) Montreal forced aligner: Trainable text–speech alignment using Kaldi. *Interspeech* 2017: 498–502. https://doi.org/10.21437/Interspeech.2017-1386

Mennen I (2004) Bi-directional interference in the intonation of Dutch speakers of Greek. *Journal of Phonetics* 32(4): 543–563. https://doi.org/10.1016/j.wocn.2004.02.002

Nagamine T (2024) Formant dynamics in second language speech: Japanese speakers' production of English liquids. *The Journal of the Acoustical Society of America* 155(1): 479–495. https://doi.org/10.1121/10.0024351

Nicenboim B, Vasishth S, Engelmann F, et al. (2018) Exploratory and confirmatory analyses in sentence processing: A case study of number interference in German. *Cognitive Science* 42: 1075–1100. https://doi.org/10.1111/cogs.12589

Nishi K, Strange W, Akahane-Yamada R, et al. (2008) Acoustic and perceptual similarity of Japanese and American English vowels. *The Journal of the Acoustical Society of America* 124(1): 576–588. https://doi.org/10.1121/1.2931949

Oh GE, Guion-Anderson S, Aoyama K, et al. (2011) A one-year longitudinal study of English and Japanese vowel production by Japanese adults and children in an English-speaking setting. *Journal of Phonetics* 39(2): 156–167. https://doi.org/10.1016/j.wocn.2011.01.002

R Core Team (2024) *R: A Language and Environment for Statistical Computing* [Computer software]. R Foundation for Statistical Computing. Available at: https://www.R-project.org (accessed June 2026).

Ramus F, Nespor M and Mehler J (1999) Correlates of linguistic rhythm in the speech signal. *Cognition* 73(3): 265–292. https://doi.org/10.1016/s0010-0277(99)00058-x

Riney TJ and Takagi N (1999) Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning* 49(2): 275–302. https://doi.org/10.1111/0023-8333.00089

Riney TJ, Takagi N, Ota K, et al. (2007) The intermediate degree of VOT in Japanese initial voiceless stops. *Journal of Phonetics* 35(3): 439–443. https://doi.org/10.1016/j.wocn.2006.01.002

Sancier ML and Fowler CA (1997) Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics* 25(4): 421–436. https://doi.org/10.1006/jpho.1997.0051

Saud Alharbi A, Foltz A, Kornder L, et al. (2023) L2 acquisition and L1 attrition of VOTs of voiceless plosives in highly proficient late bilinguals. *Second Language Research* 39(4): 1133–1163. https://doi.org/10.1177/02676583221124011

Stoehr A, Benders T, Van Hell JG, et al. (2017) Second language attainment and first language attrition: The case of VOT in immersed Dutch–German late bilinguals. *Second Language Research* 33(4): 483–518. https://doi.org/10.1177/0267658317704261

Tajima K and Port RF (2004) Speech rhythm in English and Japanese. In: Local J, Ogden R and Temple R (eds) *Phonetic Interpretation Papers in Laboratory Phonology VI*. Cambridge University Press, pp.322–339. https://doi.org/10.1017/cbo9780511486425.020

Tobin SJ, Nam H and Fowler CA (2017) Phonetic drift in Spanish–English bilinguals: Experiment and a self-organizing model. *Journal of Phonetics* 65: 45–59. https://doi.org/10.1016/j.wocn.2017.05.006

Turner J (2023) Phonetic development of an L2 vowel system and tandem drift in the L1: A residence abroad and L1 re-immersion study. *Language and Speech* 66(3): 756–785. https://doi.org/10.1177/00238309221133100

Yazawa K, Konishi T, Pérez-Ramón R, et al. (2024) More foreign-accented but more comprehensible: Attrition and amelioration of L1 speech in proficient L2 learners. *F1000Research* 13: 867. https://doi.org/10.12688/f1000research.148193.1

Yazawa K, Whang J, Escudero P, et al. (2023) Spectral and temporal implementation of Japanese speakers' English vowel categories: A corpus-based study. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 14(1): 1–33. https://doi.org/10.16995/labphon.6427

Zhang C (2022) *A guide for the R Package 'rhythm_metrics'* [Computer software]. Available at: https://osf.io/kfnzt (accessed June 2026).