Comp. by: M.Sivaraman Stage: Revises1 Chapter No.: 10 Title Name: Rao
Date:23/11/23 Time:10:16:47 Page Number: 208

10 An Individual-Differences Perspective on Variation in Heritage Mandarin Speakers

Charles B. Chang and Yao Yao

10.1 Introduction

In spite of patterns observed in adult second language (L2) learning across diverse L2s and native language (L1) groups, it has long been known that there are individual differences (IDs) in the process and outcome of L2 learning. IDs may arise for several reasons, including intrinsic variation in the cognitive and socio-affective properties of the learner (Dewaele & Furnham, 2000; Dörnyei & Skehan, 2003; Milovanov et al., 2010; Robinson, 2001; Skehan, 2002). Recent work in L2 perception and production has begun to devote systematic attention to understanding such IDs, finding predictive value for factors such as cue weighting patterns, non-linguistic perceptual abilities, the phonetics of L1 categories, L2 working memory, and neural correlates (e.g., Darcy et al., 2015; Idemaru et al., 2012; Kartushina & Frauenfelder, 2013, 2014; Perrachione et al., 2011; Raizada et al., 2010; Schertz et al., 2015, 2016).

In the study of heritage language (HL) sound systems, too, IDs have often been observed; in fact, a recurring theme of this literature has been the marked degree of inter-speaker variability among heritage speakers (HSs), which may exceed that seen in more L1-dominant native speakers and even in late-onset L2 learners (Chang, 2021). One possible contributor to this comparatively high inter-speaker variability is experiential; that is, unlike native speakers raised and educated continuously in the HL or instructed L2 learners acquiring a standard variety, HSs' lived experience with their HL, including the socio-demographic characteristics of their home environment that often provides the majority of their exposure and use opportunities related to the HL, may differ dramatically between individuals. Consequently, it is reasonable to believe that IDs in the HL sound system may arise not only from variation in cognitive and socio-affective dimensions, but also from socio-demographic differences (e.g., presence versus absence of an HL-speaking grandparent and/or sibling in the home; exposure to a non-standard variety of the HL). In the current study, our focus will be on these socio-demographic differences.

Our previous work on US-based HSs of Mandarin Chinese (Chang et al., 2009, 2010, 2011; Chang & Yao, 2016, 2019) collected extensive socio-demographic data, but ultimately used only three variables to classify participants into groups: (1) age of acquisition of Mandarin, (2) frequency of current Mandarin use, and (3) length of residence in Mandarin-speaking regions. For the purposes of that research, participants were classified as *native speakers* if they had been both born and raised in a Mandarin-speaking region; *late L2 learners* if they had been born in the US, raised in an English-speaking home (i.e., not raised in a Mandarin-speaking home), and not exposed to Mandarin until the age of 18; and *HSs* otherwise. Given the wide range of HSs' Mandarin use and residence in Mandarin-speaking regions, the HSs were further divided into two subgroups. Generally, HSs who reported using Mandarin at home more than half of the time were assigned to a *high exposure* subgroup, whereas those who reported using Mandarin at home half of the time or less were assigned to a *low exposure* subgroup, unless they had spent several years living in a Mandarin-speaking region.

Using these group designations and taking a group-based approach, our previous research produced several insights regarding Mandarin HSs. First, we found that, of the three groups, HSs were the most consistent at maintaining both language-internal and crosslinguistic contrasts between vowels, fricatives, and stop consonants (Chang et al., 2011). Second, in regard to tones, HSs were more native-like than L2 learners in some, but not all, acoustic aspects of tone production, such that their tones differed auditorily from both native speakers' and L2 learners'; in addition, HSs' tones showed the highest levels of variability, which may have played a role in the finding that HSs were the most difficult group for listeners to classify demographically (Chang & Yao, 2016). In follow-up work on Mandarin's "neutral" tone (T0), we additionally found that HSs produced T0 with different durations than the other groups (particularly in contexts where T0 is non-obligatory), leading to lower intelligibility of HSs' T0 productions compared to L2 learners' despite higher goodness ratings for HSs' intelligible T0 productions. We attributed this in part to greater variability in HSs' dialectal exposure as well as educational exposure to standard norms for Mandarin (Chang & Yao, 2019). On the basis of these results, we argued that "phonetic differences between [HSs] and [L2 learners] are not unidirectional, but instead vary across aspects of the language in accordance with differences in speakers' linguistic experience" (p. 2291).

In the current study, we take this argument further by systematically examining, using multiple factor analysis (MFA; for an overview, see Abdi et al., 2013), all of the experiential and other socio-demographic variables that we collected data on in our previous research. Although few linguistic studies have used MFA before, some have used the related techniques of principal component analysis (PCA; e.g., Bowles et al., 2016; Debras, 2017; Favier, 2020; Hodge et al., 2019) or multiple correspondence analysis (Iosad & Lamb, 2020), and MFA has been used in the neighboring field of semiotics (Piqueras-Fiszman et al., 2011). MFA is a generalization of the PCA method. Like regular PCA, MFA provides a means of discovering a set of new orthogonal variables (i.e., principal components, or "dimensions") in a high-dimensional dataset, the contribution of each original variable to the dimensions, and the representation of each individual in terms of dimensions. Unlike regular PCA, however, MFA is able to handle datasets containing both numeric and categorical variables and to balance the contribution of different groups of variables. MFA is thus ideally suited to the mixed, high-dimensional dataset examined in this study.

The current study had two main goals pertaining to the analysis of the large dataset that was compiled over the course of our previous research. For the purposes of this study, each individual speaker was given a detailed *demographic profile* (comprising more than 100 socio-demographic variables coded from their background questionnaire, some categorical and others numeric) and a detailed *phonetic profile* (comprising hundreds of numeric phonetic variables) for each language. Our primary goal was thus to demonstrate how such integrative, high-dimensional profiles could be boiled down to a low number of dimensions that still capture the majority of variance among individuals, allowing for a more manageable exploration of IDs and of the relative development of different types of phonetic variables. A secondary goal was to examine how the group classifications used in our previous work, which were based on only a few selected variables, would compare with the results of MFA.

Crucially, we take the MFA approach here not only because it matches well with our data, but also because this approach, which takes a much broader look at the data than was taken in any of our previous studies, offers several advantages. First, it allows us to identify speaker groups more confidently, on the basis of a wide swath of socio-demographic variables rather than a few selected ones. Second, it produces more nuanced visualizations of IDs in HSs' sound systems, which incorporate multiple phonetic variables. And third, it permits us to relate these IDs to socio-demographic factors at a higher level (i.e., the level of orthogonal "dimensions"). In the rest of this chapter, we provide, as proof of concept, a brief report on our MFA findings, concluding with thoughts on the joint investigation of IDs in the HL and the majority language and on the pros and cons of group-based and individual-centered approaches.

10.2 Methodology

10.2.1 Participants

The participants reported on in Chang et al. (2009, 2010, 2011) and Chang and Yao (2016, 2019) comprised three groups of Mandarin speakers living in the

US at the time of study: native speakers born and raised in a Mandarin-speaking region (the NM group; *n* = 6), late-onset L2 learners from an L1 English background (the L2 group; *n* = 5), and HSs (the HL group in past work; *n* = 15). To streamline labeling in graphs pertaining to individual differences, we enumerate members of these groups respectively with the prefixes N, L, and H (i.e., the same labels shown in Chang et al., 2011, p. 3977).

Because the HSs spanned a wide range of experience with Mandarin, in previous work, this group was further divided into two subgroups. Nine speakers comprised a *high exposure* (HE) subgroup and are labeled in visualizations with *a* (i.e., Ha7–Ha15), while six others comprised a *low exposure* (LE) subgroup and are labeled with *b* (i.e., Hb16–Hb21). Additional background information about the HE and LE subgroups, as well as the NM and L2 groups, is provided in Chang et al. (2011, pp. 3967–3968, 3977), and we discuss the socio-demographic differences among these participants in more detail below.

10.2.2 Procedure and Materials

In one session, participants completed an elicited production task and a detailed language background questionnaire that asked about early exposure to Mandarin and other languages, the language backgrounds of immediate and extended family, educational experience with Mandarin, and perceived proficiency levels.[1] The production task was completed in a sound-attenuated booth in two blocks (i.e., one block for Mandarin and one block for English), with a break between blocks. In each block, participants were shown a target item on a flashcard and asked to utter the item aloud, on its own. The experimenter cycled through the flashcards four times (shuffling the deck after each cycle), resulting in four tokens of each item.

Target items comprised a total of fifty-nine Mandarin words and phrases and thirty-two English words, selected to elicit productions of the back rounded vowels /u o(u)/ in Mandarin and English, the front rounded vowel /y/ in Mandarin, the stops in Mandarin (/p t k p^h t^h k^h/) and English (/b d g p t k/), the sibilant fricatives in Mandarin (/s ʂ ɕ/) and English (/s ʃ/), and the four lexical tones (T1–T4) and "neutral" tone (T0) of Mandarin. In short, the materials included Mandarin sounds that resemble English counterparts to varying degrees, as well as Mandarin-specific features.

Further details about the testing procedure, as well as the full lists of target items, are provided in Chang et al. (2011, pp. 3968–3969, 3978) and Chang and Yao (2016, pp. 139–140).

[1] The full questionnaire is available at https://osf.io/u9wz4/.

10.2.3 Socio-Demographic and Acoustic Data

A total of 133 socio-demographic/language background variables were coded from questionnaire responses as either categorical or numeric (i.e., continuous), depending on the relevant questionnaire item. For example, the variable of *age of arrival in the US* was numeric, while the variable of *whether or not the participant spoke primarily Mandarin to their mother* was categorical. Because not all participants completed the questionnaire in its entirety, analyses of the full participant sample focused on a subset of nine variables that contained data for all or most of the sample; analyses of just the HSs focused on a subset of thirty-seven variables (discussed further in Section 10.3).

A total of 110 acoustic variables (e.g., F_1 of Mandarin /u/ in an alveolar context; voice onset time (VOT) of English /t/) were measured via acoustic analyses in Praat (Boersma & Weenink, 2016), which are described in detail in Chang et al. (2011, p. 3969), Chang and Yao (2016, pp. 141–142), and Chang and Yao (2019, p. 2293). Acoustic variables comprised the first three vowel formants (F_1, F_2, F_3), VOT following prevocalic stops, spectral components of fricatives (specifically, the peak amplitude frequency and centroid), and fundamental frequency (f_0) and durational properties of tones (e.g., the “turning point” of T3).[2]

10.2.4 Statistical Analysis

To reduce the dimensionality of the socio-demographic and acoustic datasets, MFA was carried out using the MFA() function in the FactoMineR package (Lê et al., 2008) in R (R Development Core Team, 2021). Related visualizations were created with the factoextra package (Kassambara & Mundt, 2020).[3] Both socio-demographic variables and acoustic variables were subjected to MFA, with all numeric variables transformed to z-scores (mean = 0, variance = 1) before being entered. A *TalkerGroup* variable group (comprising participants’ original group assignments in Chang et al., 2011; i.e., NM, L2, HE, LE) is included in the factor maps below but acted as a “latent” variable, meaning it did not contribute to the dimensions that emerged from MFA.

10.3 Results

We report here the results of four MFAs. The four MFAs are of socio-demographic variables over the full participant sample (MFA1), of

[2] The full dataset of both socio-demographic and acoustic variables is available at https://osf.io/9rehm/.

[3] High-resolution color versions of all visualizations in this chapter can be viewed at https://osf.io/5hyjf/.

socio-demographic variables over HSs only (MFA2), of phonetic variables in Mandarin production (MFA3), and of phonetic variables in English production (MFA4). The results of all MFAs indicated that the first two dimensions (i.e., the two most important components: Dim1, Dim2) together accounted for the plurality, and usually the majority, of variance (i.e., 40–67%), so we focus on these two dimensions in the discussion that follows. Note that we are most interested in how the variables, variable groups, and individuals map onto the dimensions emerging from MFA. As in regular PCA, it is not straightforward to identify a precise meaning for each dimension, other than knowing that the dimensions comprise an orthogonal space that optimally captures the variance of the dataset. Thus, we focus our discussion on the distribution of variables, variable groups, and individuals in this space, which provided insight into the internal structure of the dataset.

We begin by presenting the results of MFA1, which helped us understand how participant (i.e., talker) groups differed from each other socio-demographically. This analysis focused on the subset of nine variables about which questionnaire data were available for all participants, which were sorted into three variable groups.[4] The first variable group (*AoAr*) comprised the only numeric variable, age of arrival in the US (AoAr). The second (*SelfLg*) comprised self-reported L1 (Chinese, non-Chinese, or both Mandarin and English), country of birth (CoB; Chinese-speaking or non-Chinese-speaking), and best language (BestLg; Chinese, English, or both Mandarin and English). The third (*ParentsLg*) comprised home Mandarin variety in terms of a Mainland China versus Taiwan/Singapore dichotomy (ManVrty1) and in terms of a Northern versus Southern (e.g., Shanghai, Chaozhou, Guangzhou, Taiwan) dichotomy (ManVrty2), whether or not a participant's mother (MomL1) and father (DadL1) spoke some variety of Chinese as their L1, and the primary language(s) spoken at home (HomeLg; Chinese, non-Chinese, or both Mandarin and English). The variables ManVrty1 and ManVrty2 were coded according to what a participant reported about their family's, in particular their parents', variety of Mandarin; L2 learners were thus coded as 'NA' for both of these variables.

The factor maps in Figure 10.1 show the correlations between each variable (group) and the first two dimensions. The coordinate of a given variable (group) on a given dimension indicates how much the specific variable (group) contributes to the given dimension. Similarly, the maps in Figure 10.2 show how each individual is represented by the first two dimensions. More specifically, Figure 10.1a shows that the three variable groups (*AoAr, SelfLg,*

4 For reasons of space, we do not generally provide an exhaustive list of variables within each variable group in the text. However, for MFA1–MFA4, an exhaustive list of variables is provided in supplementary materials available at https://osf.io/5hyjf/.

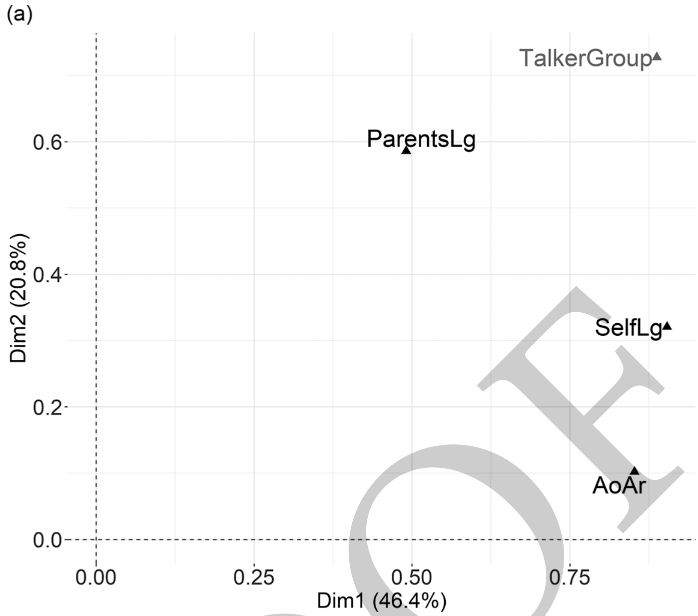

Figure 10.1 Variable groups and categorical variables (by level) in MFA1. (a) Representations of variable groups, including latent *TalkerGroup*, in terms of the first two dimensions (Dim1, Dim2) of MFA1 over the whole participant sample; (b) Mean representations of each level of the categorical variables (prefixes indicate variable group; Self = *SelfLg*, Prnt = *ParentsLg*).

ParentsLg) contributed to Dim1 and Dim2 to different degrees, while both dimensions were relevant for the participant groupings encompassed in *TalkerGroup*. Figure 10.1b plots each variable against Dim1 and Dim2, and reveals a triangular formation of the categorical variables that is also evident in the orientation of individuals in Figure 10.2a; specifically, the upper-left corner of the Dim1–Dim2 space contained mostly "non-Chinese" features (in Figure 10.1b) and L2 participants (in Figure 10.2a), while the upper-right corner contained mostly "Chinese" features and NM participants, and the lower-middle corner contained a mix of "Chinese," "English," and "Mandarin+English" features and mainly HSs. Notably, Figure 10.2a was, overall, consistent with the original participant groups in Chang et al. (2011), providing evidence of the validity of those groups despite the fact that they were based on only a few demographic factors.

Example demographic profiles of three peripheral (i.e., further from the origin) members of the NM, L2, and HS groups are depicted in Figure 10.2b. These profiles show participants L25 and Ha9 converging on

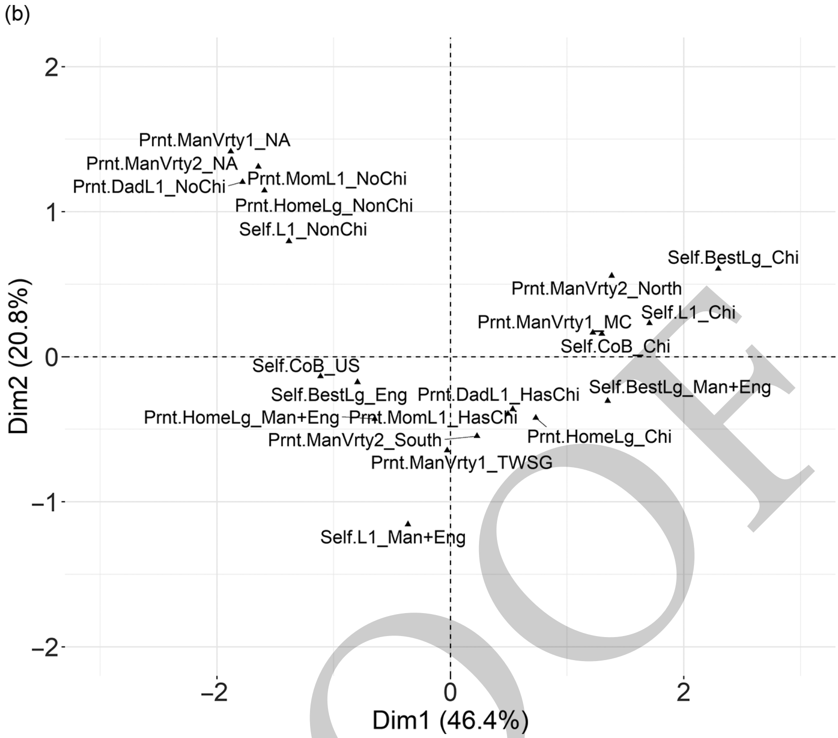

Figure 10.1 (*cont.*)

AoAr in the lower-left quadrant, but N4 projecting on *AoAr* into the upper-right quadrant, reflecting the fact that N4 had a much later AoAr than both L25 and Ha9. As for *SelfLg* and *ParentsLg*, there is more of a triangular pattern in the projections for these variable groups (i.e., distinct projections across the three individuals) although N4 and Ha9 projected into the same lower-right quadrant for *ParentsLg*, reflecting the overlap in features in this variable group for these two speakers. Thus, these demographic profiles provide a means of visualizing both similarities and differences among individuals with multifaceted backgrounds.

Moving on to MFA2, we conducted this analysis to better understand how HSs differed from each other socio-demographically. Three HSs (Ha7, Ha8, Hb20) were dropped from this analysis because they provided no data regarding Mandarin contact or their parents' AoAr; thus, MFA2 focused on twelve of the fifteen HSs. Apart from the variables included in MFA1, MFA2 included many additional variables, all those that were widely represented in, and showed variation among, the HSs. This was a set of thirty-seven variables,

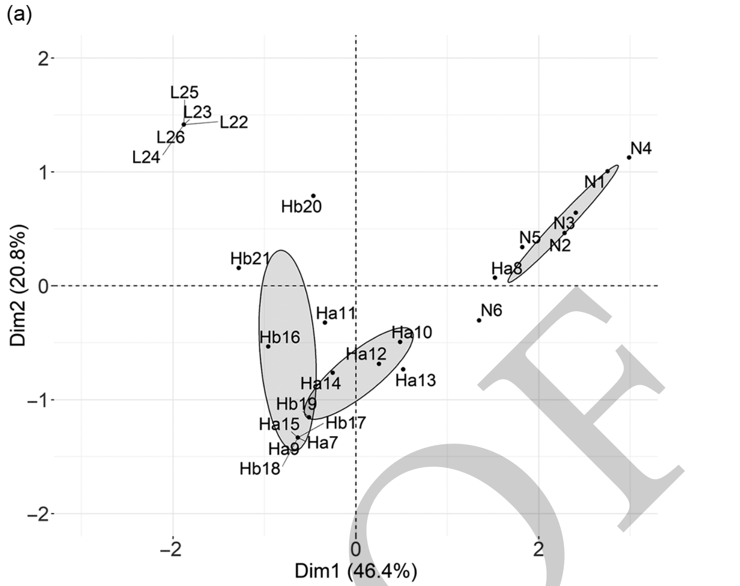

Figure 10.2 Representations of individuals according to demographic variables in MFA1. (a) Representations of individuals in terms of summed projections of all the variables in their demographic profile on Dim1 and Dim2 in MFA1 (ellipses mark the 95% CI around group means); (b) Demographic profiles of three peripheral individuals (N4, L25, Ha9) in terms of the three variable groups.

which were classified into eight variable groups. The first variable group (*ResidLen*) comprised each participant's AoAr (AoAr), length of residence in a Mandarin-speaking (YrsMan) and English-speaking environment (YrsEng), their father's AoAr (DadAoAr), and their mother's AoAr (MomAoAr). The second variable group (*SelfLg*) again comprised each participant's L1, country of birth (CoB), and best language (BestLg). The third variable group (*ParentsLg*) comprised, as above, ManVrty1, ManVrty2, DadL1, and HomeLg, as well as the primary language each participant heard from their mother and father (respectively, LgFrMom, LgFrDad; Mandarin or English) and spoke to their mother and father (respectively, LgToDad, LgToMom; Mandarin or English). The fourth variable group (*TalkFreqWParents*) comprised each participant's frequency of speaking Mandarin with their mother (Mom) and father (Dad). The fifth variable group (*Siblings*) comprised each participant's number of younger siblings (CntYng) and older siblings (CntOld), and the frequency of siblings' speaking in Mandarin to each participant (ManInput) and to their parents

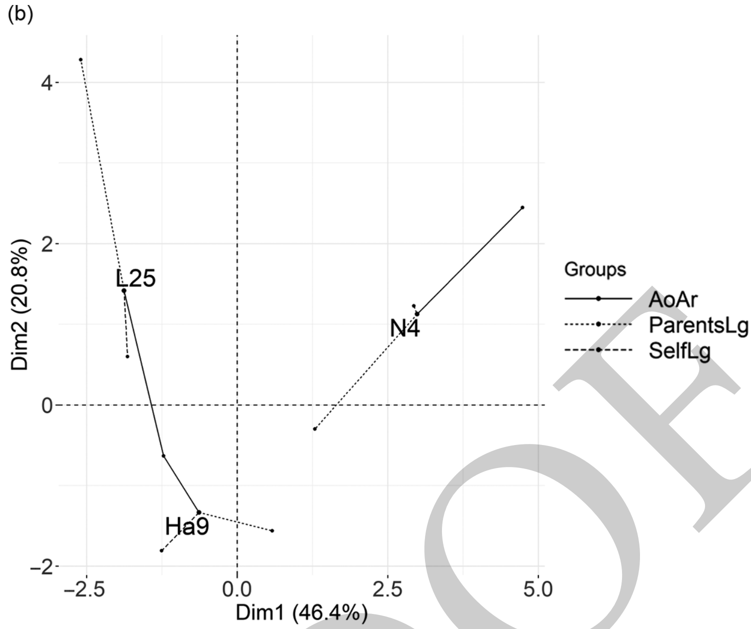

Figure 10.2 (*cont.*)

(ManToPrnt). The sixth variable group (*Grandparents*) comprised the length of residence with grandparents (YrsRes) and the frequency of grandparents' speaking in Mandarin to each participant's parents (ManToPrnt). The seventh variable group (*ManExperience*) comprised the total amount of Mandarin overhearing experience (Overhear), listening experience (Listen; i.e., Mandarin speech directed at the participant), and speaking experience (Speak), summing over several different contexts. The eighth variable group (*ManProficiency*) comprised self-ratings of Mandarin proficiency in eleven contexts: telling a story (Story), ordering in a restaurant (Order), shopping (Shop), conversing with relatives about casual topics (Converse), speaking with strangers at a community meeting (CommMtg), talking about school or work (SchoolWork), discussing politics (Pol), giving a speech (Speech), being interviewed for a job (Interview), understanding TV programs (CompTv), and understanding formal speech (CompFml).

Factor maps resulting from MFA2 are shown in Figures 10.3 and 10.4. As seen in Figure 10.3a, *TalkFreqWParents* projected almost entirely onto Dim2, whereas *ManExperience*, *ParentsLg*, *Grandparents*, and *ManProficiency* projected almost entirely onto Dim1. The remaining variable groups (*Siblings*, *ResidLen*, *SelfLg*) projected onto both dimensions. Notably,

(a)

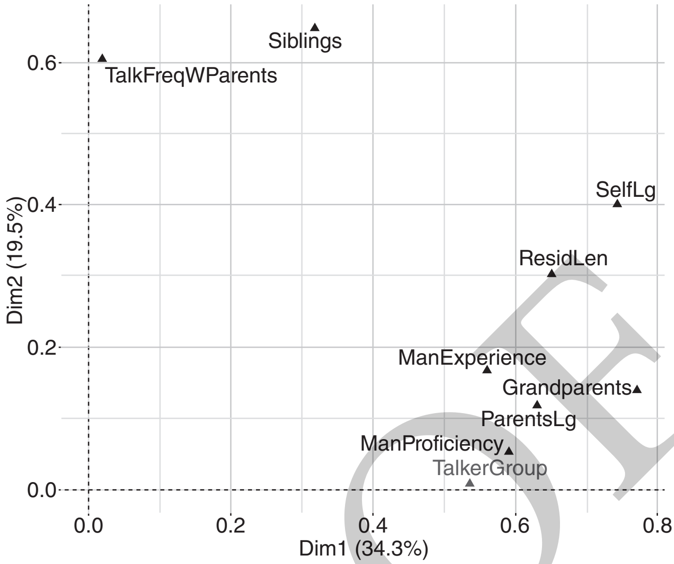

Figure 10.3 Variable groups, categorical variables (by level), and numeric variables in MFA2. (a) Representations of variable groups, including latent *TalkerGroup*, in terms of the first two dimensions (Dim1, Dim2) of MFA2 over HSs; (b) Mean representations of the categorical variables (prefixes indicate variable group; Self = *SelfLg*, Prnt = *ParentsLg*); (c) Mean representations of the numeric variables (TalkFreq = *TalkFreqWParents*, Sibs = *Siblings*, Gramps = *Grandparents*, Exp = *ManExperience*, Prof = *ManProficiency*).

TalkerGroup projected onto Dim1 only, suggesting that it is orthogonal to *TalkFreqWParents*. Figure 10.3b plots the projections of categorical variables, which showed that "Chinese" or "Mandarin" features mostly projected onto the right half of the plot and "English" features mostly onto the left half. As for the projections of numeric variables, as seen in Figure 10.3c, *ManExperience* and *ManProficiency* variables mostly projected positively onto Dim1, while *TalkFreqWParents* variables projected almost entirely onto Dim2 (positively). The *Siblings* variables CntYng and CntOld projected disparately onto Dim1, the former negatively and the latter positively. Like CntOld, the *Grandparents* variable YrsRes, along with the *ResidLen* variables AoAr, DadAoAr, MomAoAr, and YrsMan, projected positively onto Dim1. On the other hand, YrsEng projected negatively onto Dim1.

The orientation of individual HSs in the factor map is shown in Figure 10.4a, and example demographic profiles of two peripheral HSs are

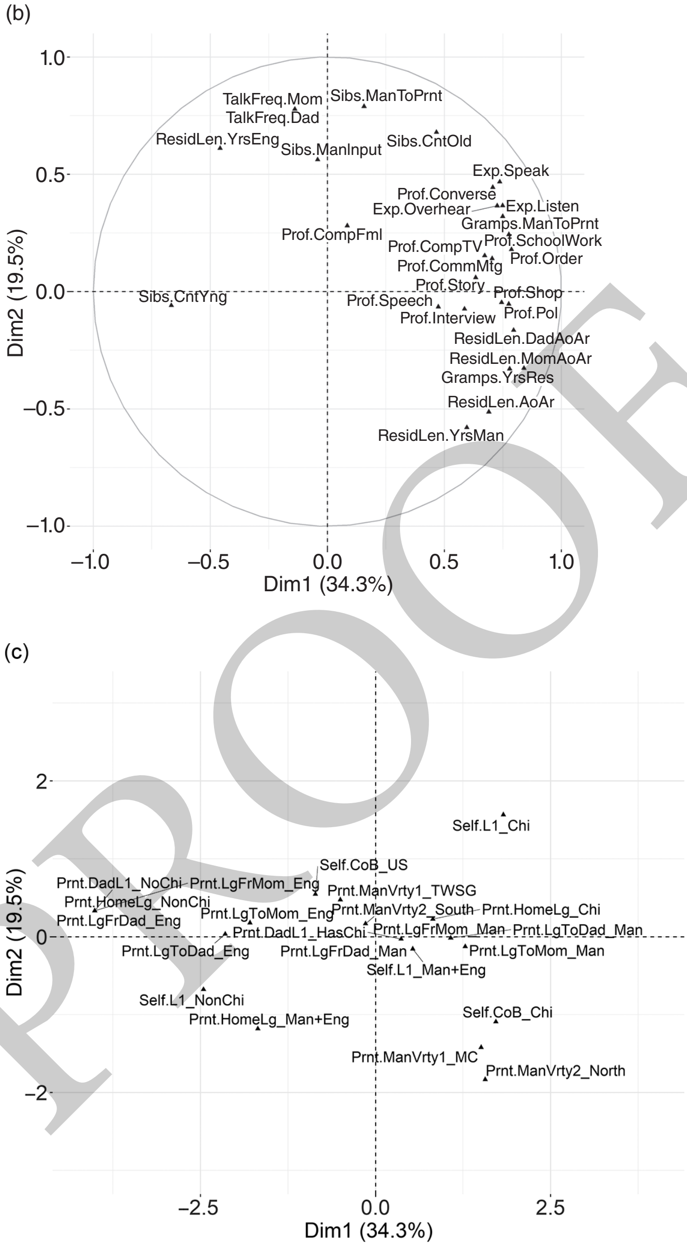

Figure 10.3 (*cont.*)

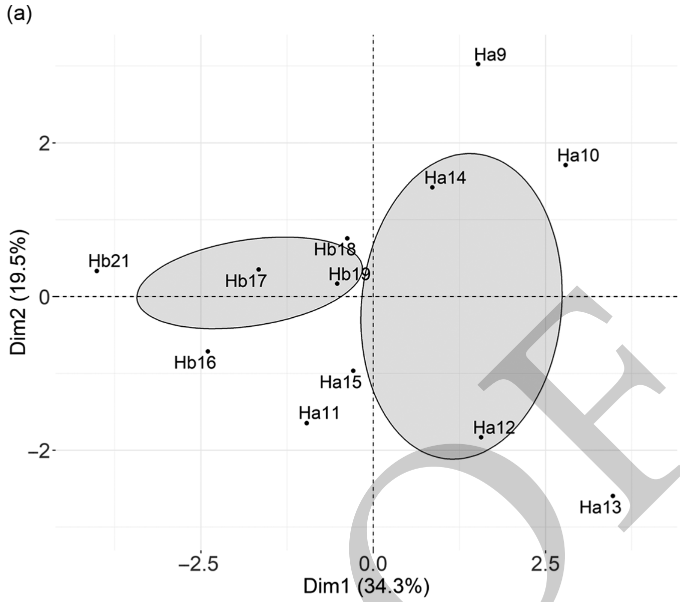

Figure 10.4 Representations of heritage speaker individuals according to demographic variables in MFA2. (a) Representations of HSs (*Ha* = HE subgroup, *Hb* = LE subgroup) in terms of summed projections of all the variables in their demographic profile on Dim1 and Dim2 in MFA2 (ellipses mark the 95% CI around subgroup means); (b) Demographic profiles of two peripheral HSs (Ha10, Hb21) in terms of the eight variable groups.

displayed in Figure 10.4b. As seen in Figure 10.4a, members of the HE subgroup (labeled *Ha*) and LE subgroup (labeled *Hb*) were distinguished mostly by Dim1, with HE speakers on the right and LE speakers on the left. Combined with Figure 10.3, this result thus suggested that the HE versus LE subgroup distinction corresponded primarily to differences in the *ManExperience*, *ManProficiency*, *ParentsLg*, and *Grandparents* variables, as opposed to *TalkFreqWParents*. As for *Siblings*, the numbers of older (CntOld) versus younger (CntYng) siblings were aligned with the HE versus LE subgroup distinction. Indeed, in Figure 10.4b, we see that participants Ha10 and Hb21 converged on *TalkFreqWParents*, but differed with respect to all the other variable groups; for Ha10, these projected onto Dim1 positively, whereas for Hb2, they did so negatively.

With regard to MFA3 and MFA4, recall that these analyses were meant to reduce the dimensionality of the phonetic variables, with MFA3 focusing on Mandarin and MFA4 on English. To confirm the validity of the results, given

(b)

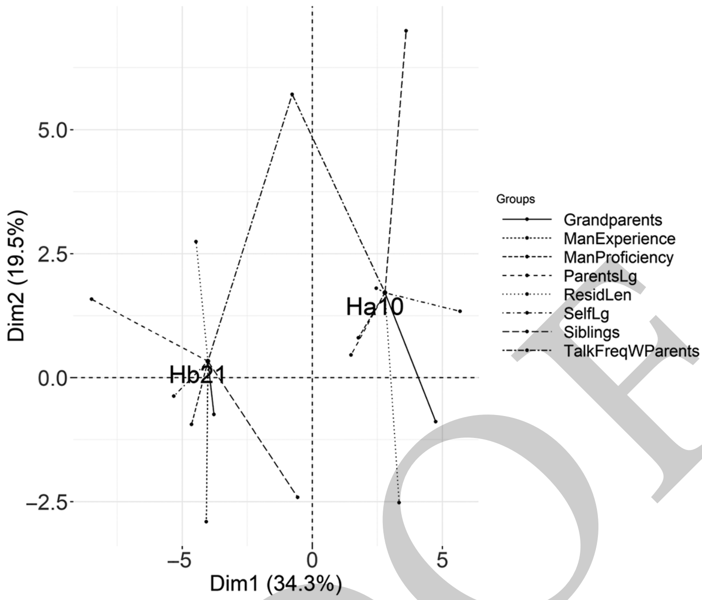

Figure 10.4 (*cont.*)

the limited size of the participant sample, we conducted cross-validation tests for both MFA3 and MFA4 by drawing twenty random samples of talkers (10 with 16 talkers, or about 60 percent of the talkers; and 10 with 21 talkers, or about 80 percent of the talkers) and replicating the MFA analyses on the acoustic data of each random sample of talkers.[5] Our report on the MFA3 and MFA4 results based on the full datasets thus includes comments about the robustness of the findings based on cross-validation using partial datasets.

Starting with MFA3, we included six variable groups in this analysis: *VowelF1* (including, e.g., u(blb), u(alv), u(plt), u(vlr), and u(glt), referring to the vowel /u/ in a bilabial, alveolar, palatal, velar, and glottal context, respectively; ou(nil), referring to /ou/ in an onsetless context, etc.), *VowelF2*, *VowelF3*, *Fricative* (e.g., ʂ.Cntrd and ʂ.PAF, referring to the centroid and peak amplitude frequency of retroflex /ʂ/), *VOT* (e.g., t and th, referring to the mean VOT of short-lag /t/ and long-lag /tʰ/, respectively, etc.), and *Tone* (e.g., T1Start, T1End, T1Range, T1Mean, T1Dur1, and T1Dur2, referring to T1's starting f_0, ending f_0, f_0 range, mean f_0, duration in isolation, and duration

[5] Factor maps from these cross-validation tests can be viewed at https://osf.io/5hyjf/.

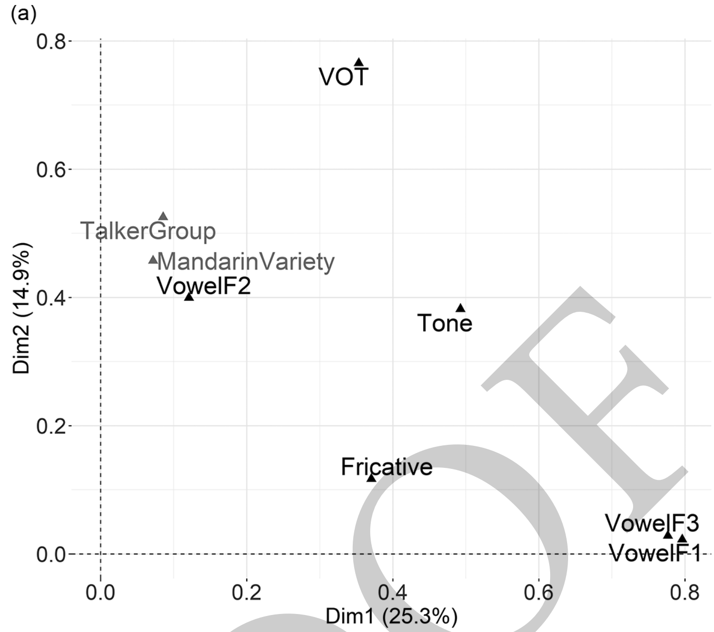

Figure 10.5 Variable groups and Mandarin phonetic (segmental and tonal) variables in MFA3. (a) Representations of variable groups, including latent *TalkerGroup* and *MandarinVariety*, in terms of the first two dimensions (Dim1, Dim2) of MFA3; (b) Mean representations of the segmental variables (prefixes indicate variable group; F1 = *VowelF1*, F2 = *VowelF2*, F3 = *VowelF3*, Fric = *Fricative*); (c) Mean representations of the tonal variables.

in context, respectively; T3TurnPt, T3HalfT3Rate1, and T3HalfT3Rate2 (Final), referring to T3's turning point and rate of realization as "half Tone 3" in isolation and in the final syllable of a multisyllabic context; ToDur2 and ToDur2(NonObl), referring to T0's duration in context overall and in non-obligatory contexts specifically, etc.).

Factor maps resulting from MFA3 are shown in Figures 10.5 and 10.6. As seen in Figure 10.5a, *Fricative*, *VowelF1*, and *VowelF3* projected mostly onto Dim1 and *VowelF2* onto Dim2, while *VOT* and *Tone* projected onto both Dim1 and Dim2. In addition to *VowelF2*, both *TalkerGroup* and *MandarinVariety* projected mostly onto Dim2 (close to each other, because they were highly correlated; e.g., L2 learners were coded as 'NA' for *MandarinVariety*, as described above), suggesting that *TalkerGroup* was distinguished mostly by *VowelF2*, as well as *VOT* and *Tone*. Cross-validation results showed these patterns to be quite consistent, although there was some

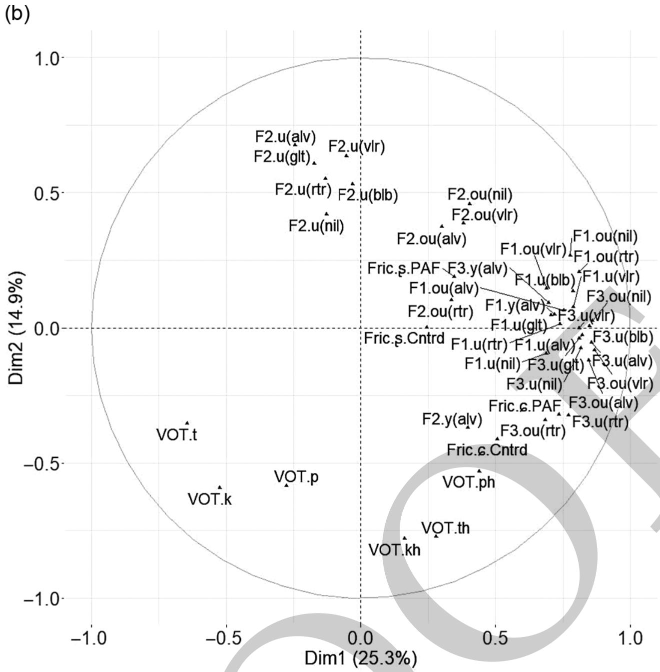

Figure 10.5 (*cont.*)

variation for *Fricative* (which projected more onto Dim2 in some tests). Figures 10.5b and 10.5c plot the projections of variables in the segmental and tonal variable groups, respectively. Converging with Figure 10.5a, these figures indicate that variables in the *Fricative*, *VowelF1*, and *VowelF3* variable groups projected mostly onto Dim1, while those in the *VowelF2*, as well as the *VOT*, variable group projected mostly onto Dim2. In particular, *VowelF2* variables projected onto Dim2 positively and *VOT* variables did so negatively. On the other hand, *Tone* variables occupied a relatively wide area, with some (e.g., T1Dur2, T3Dur2) projecting positively onto Dim2 and others (e.g., T3HalfT3Rate2(NonFinal), T4Dur1) doing so negatively. Cross-validation showed variation for *Fricative* variables (which, as above, projected heavily onto Dim2 in several tests), as well as inconsistency in the direction of projections; in some tests, *VowelF2* variables projected onto Dim2 negatively and *VOT* variables did so positively. There was also some inconsistency in the projections of T1Dur2 and T3Dur2, on the one hand, and T3HalfT3Rate2 (NonFinal) and T4DUR1, on the other; although the two pairs of *Tone* variables consistently projected in different directions onto Dim2, the specific direction of their projection differed across tests. Change in the orientation of

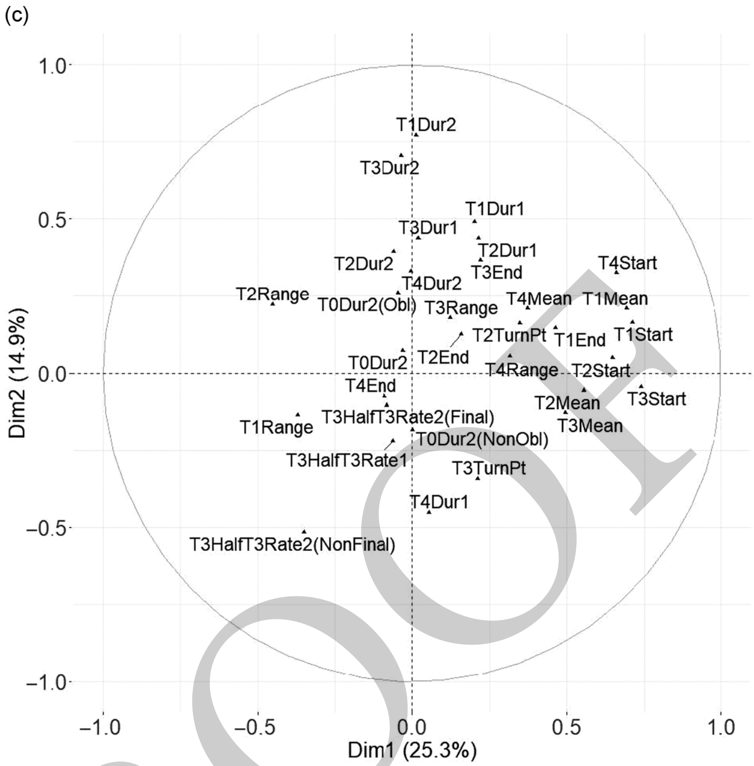

Figure 10.5 (*cont.*)

Tone variables on Dim2, however, was always in sync with change in the orientation of *VowelF2* and *VOT* on Dim2.

The orientation of individuals in the factor map of Mandarin production is shown in Figure 10.6a, and Mandarin phonetic profiles of four peripheral individuals are provided in Figure 10.6b. As seen in Figure 10.6b, participants L23 and Ha12 had the highest projections of *VowelF2* onto Dim2, followed by Hb20, and then N5. With respect to *VOT* as well, participant L23 had the highest projection, followed by Hb20, and then Ha12 and N5. In conjunction with Figures 10.5b and 10.5c, these results suggest that, overall, L2 participants tended to have higher F_2 values and lower VOTs than NM participants, with HSs patterning in between these two groups. Cross-validation showed this tendency to be relatively consistent, but in some tests the orientation of talker groups (see Figure 10.6a) was in the reverse direction along Dim2 (i.e., NM > HS > L2), which was correlated with a reversed orientation of the *VowelF1*, *VOT*, and *Tone* variables along Dim2. Additionally, the HS and NM groups

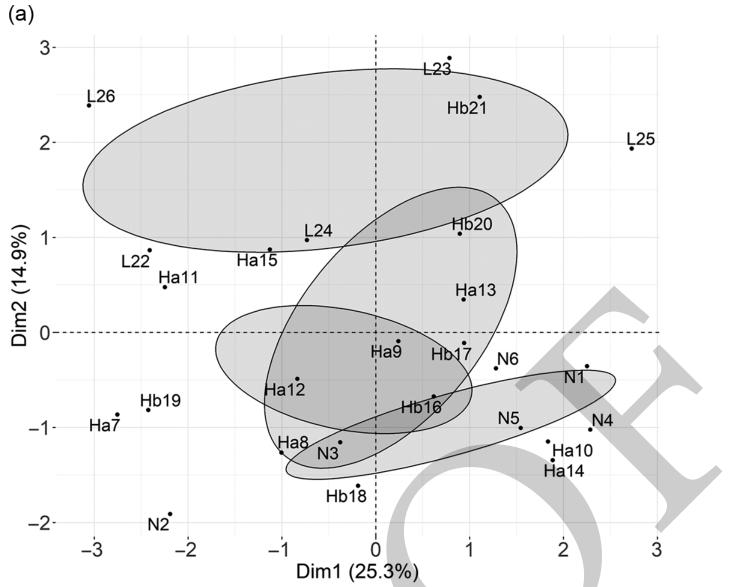

Figure 10.6 Representations of individuals according to Mandarin phonetic variables in MFA3. (a) Representations of individuals in terms of summed projections of all the variables in their Mandarin phonetic profile on Dim1 and Dim2 in MFA3 (ellipses mark the 95% CI around group means); (b) Mandarin phonetic profiles of four peripheral individuals (N5, L23, Ha12, Hb20) in terms of the six variable groups.

often appeared more merged than they appear in Figure 10.6a, likely due to the smaller sample of HS/NM talkers drawn for the cross-validation tests.

Given the different Mandarin dialectal backgrounds found among the HS and NM participants, we conducted additional analyses examining the role of *MandarinVariety*, as shown in Figure 10.7. In Figure 10.7a, which plots individuals separated by ManVrty1 on the left and ManVrty2 on the right, it is clear in the left panel that participants whose home variety was Taiwan or Singapore Mandarin (TWSG; in black) and those whose home variety was Mainland Chinese (MC; in gray) were highly overlapped, yet slightly shifted apart on Dim1; in the right panel, a similar pattern is evident for speakers whose home variety was Southern (black) versus Northern (gray). These results thus suggest that the TWSG versus MC contrast resembled the Southern versus Northern contrast, both corresponding mostly to Dim1; however, although this means that these dialectal contrasts are likely to be located in variables from the *VowelF1*, *VowelF3*, and/or *Fricative* variable groups (consistent with the individual phonetic profiles in Figure 10.7b), it remains

(b)

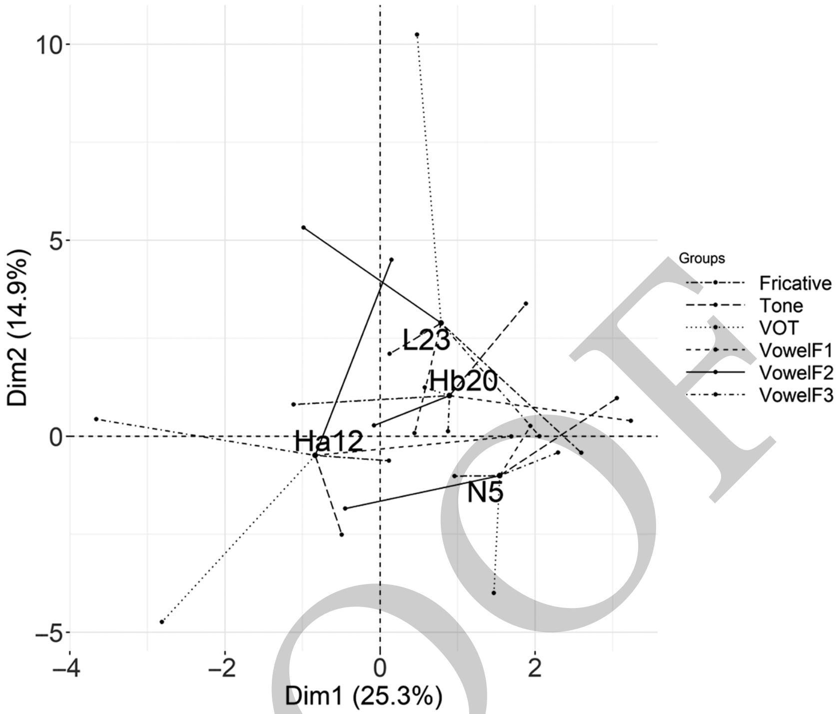

Figure 10.6 (*cont.*)

unclear exactly which of these variables contribute to the dialectal contrasts. Cross-validation showed a consistent parallelism of the TWSG versus MC and Southern versus Northern contrasts along Dim1, although in a small number of tests (mostly with 60% talker samples), these contrasts were not clearly visible, and the talker groups were more merged, again likely due to the smaller sample of talkers.

Finally, MFA4, focusing on English production, included five variable groups (the same ones in MFA3 except for *Tone*), and the factor maps resulting from this analysis are shown in Figures 10.7 and 10.8. As seen in Figure 10.7a, *Fricative*, *VowelF1*, and *VowelF3* projected mostly onto Dim1, and *VowelF2* and *VOT* did so mostly onto Dim2 (cf. Figure 10.5a). Furthermore, *TalkerGroup* projected mostly onto Dim2, suggesting that *TalkerGroup* was distinguished mostly by *VowelF2* and *VOT*. These projection patterns remained consistent in cross-validation. Consistent with Figure 10.7a, Figure 10.7b indicates that variables in the *Fricative*, *VowelF1*, and *VowelF3* variable groups mostly projected onto Dim1, and those in the

(a)

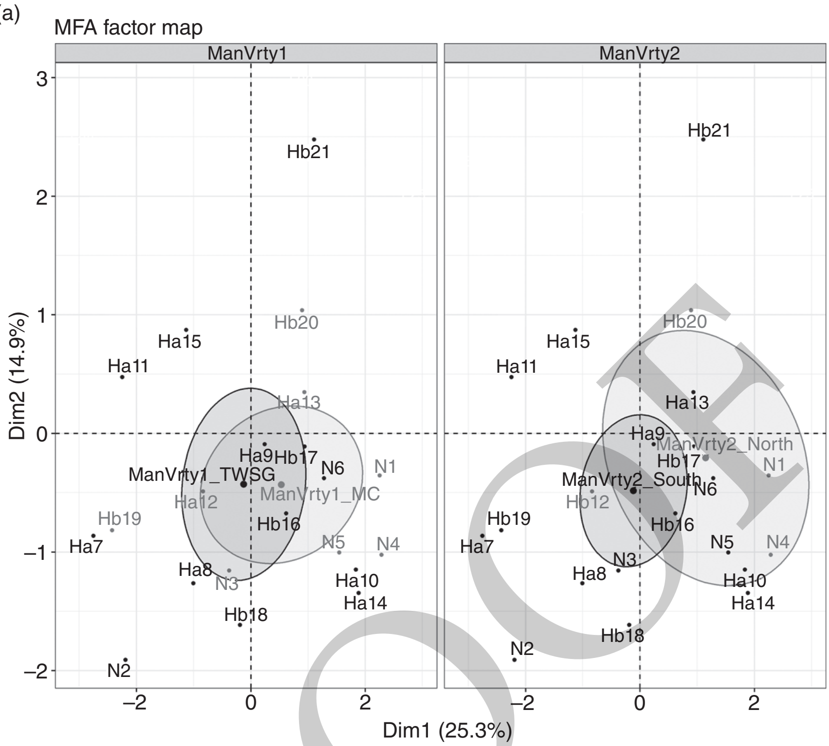

Figure 10.7 Representations of heritage speaker and native Mandarin individuals by Mandarin variety in MFA3. (a) Representations of HS and NM participants, along with *MandarinVariety* variables (TWSG = Taiwan/Singapore, MC = Mainland China), on Dim1 and Dim2 in MFA3 (ellipses mark the 95% CI around subgroup means); (b) Mandarin phonetic profiles of four individuals (N2, N4, Ha11, Ha12) contrasting in *TalkerGroup* and *MandarinVariety*.

VowelF2 and *VOT* variable groups mostly did so onto Dim2. However, while all *VowelF2* variables projected onto Dim2 positively, *VOT* variables showed a disparity in terms of their projections onto Dim2; some (T, K, P, D) projected positively, while others (G, B) projected negatively, with smaller coordinates. Cross-validation showed some variation in the projections of the *Fricative* and *VOT* variables, but an overall clear trend for *Fricative* variables to project mostly onto Dim1 and *VOT* variables onto Dim2, in line with the overall pattern. Further, while in some tests the *VOT* variables all projected positively onto Dim2, in general, the T, K, P, and D variables appeared higher (i.e., more positive) than the G and B variables, in line with the overall pattern.

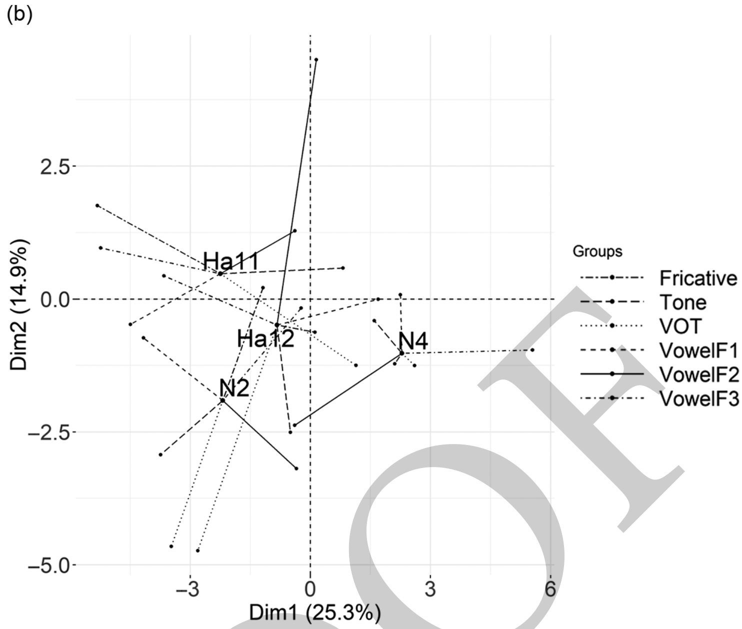

Figure 10.7 (*cont.*)

The orientation of individuals in the factor map of English production is shown in Figure 10.9a, and English phonetic profiles of four peripheral individuals are given in Figure 10.9b. Consistent with Figure 10.8a, Figure 10.9a shows that participant groups were distinguished mostly along Dim2, with LE (*Hb*) participants at the top of Dim2, followed by L2 and HE (*Ha*) participants (who largely overlapped), and then by NM participants at the bottom. This pattern was consistent across most of the cross-validation tests. Again, since Dim2 comprised mostly *VowelF2* and *VOT*, this suggests that participant groups were primarily distinguished by variables in these two variable groups. Indeed, the phonetic profiles in Figure 10.9b show that the individuals representing the four participant groups clearly differed in *VowelF2* variables (solid line). Given the coordinates of *VowelF2* variables in Figure 10.8b (i.e., projecting positively onto Dim2), the profiles indicate that participants Hb18 and L24 both had higher F_2 measures, followed by Ha15, and then N3, suggesting that Hb18 and L24 produced more fronted (therefore, more canonically English-like) back vowels than Ha15 and N3. As for *VOT* variables (dotted line), participants N3, Ha15, and L24 had overall similar *VOT* coordinates on the Dim1 $\times$ Dim2 plane (in the lower-left quadrant), whereas Hb18's *VOT* coordinates were in the upper-right quadrant, suggesting that

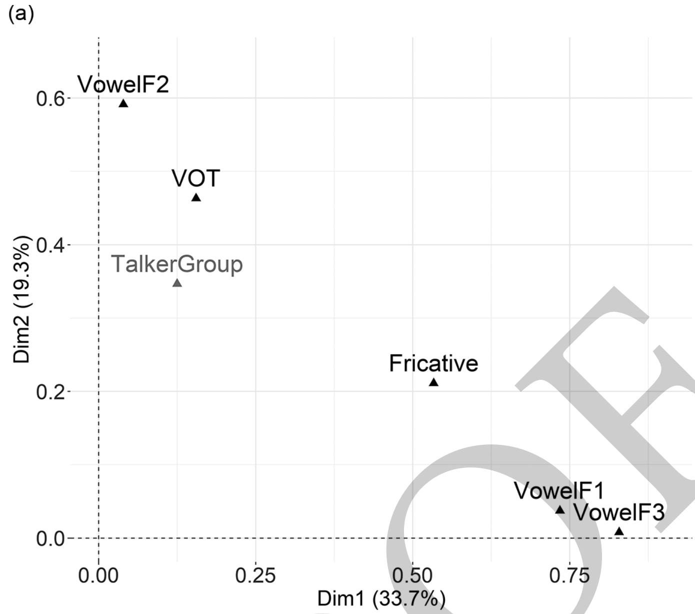

Figure 10.8 Variable groups and English phonetic variables in MFA4. (a) Representations of variable groups, including latent *TalkerGroup*, in terms of the first two dimensions (Dim1, Dim2) of MFA4; (b) Mean representations of phonetic variables (prefixes indicate variable group; F1 = *VowelF1*, F2 = *VowelF2*, F3 = *VowelF3*, Fric = *Fricative*).

Hb18 produced longer VOTs for the (mostly) voiceless stops projecting onto Dim2 positively (T, K, P, D), and thus larger VOT distinctions as compared to the voiced stops projecting onto Dim2 negatively (G, B).

10.4 Discussion and Conclusions

Returning to our study goals, we set out to test an approach to reducing the dimensionality of detailed demographic and phonetic production profiles that could support the investigation of IDs in production as well as the exploration of how different phonetic aspects of a given language may develop in relation to each other. This approach, based on MFA, produced three main insights. First, the overall organization of participants in MFA of a broader set of socio-demographic variables was generally consistent with the group divisions used in our previous work (Chang et al., 2011; Chang & Yao, 2016, 2019), providing evidence of the validity of the original groups that were based on only a few variables. At the same time, however, the few cases of ambiguous

(b)

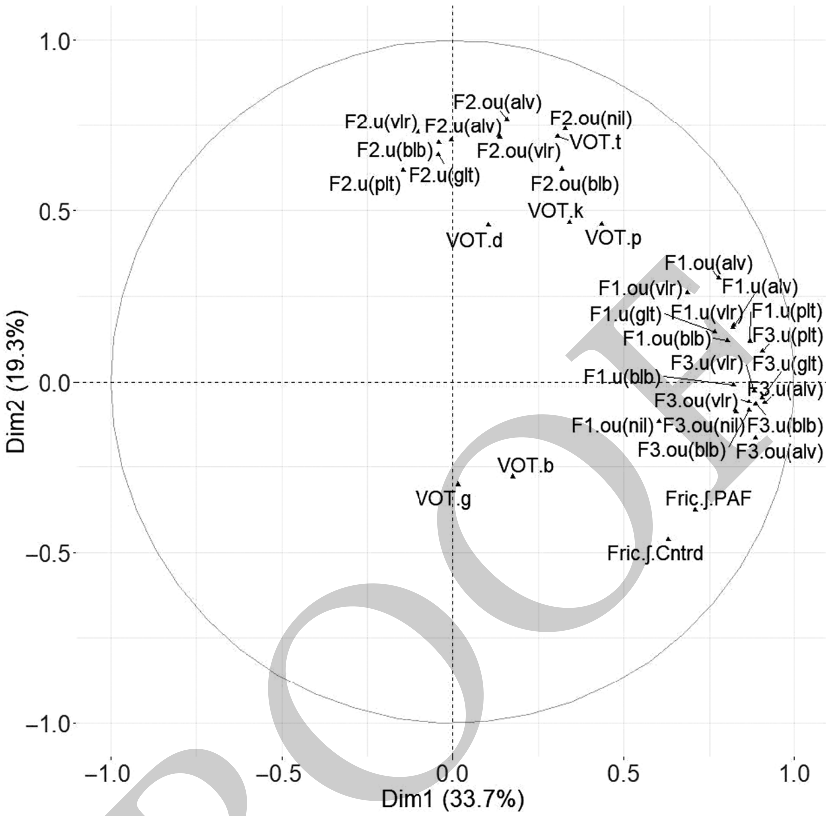

Figure 10.8 (*cont.*)

or apparently erroneous group affiliation within the two-dimensional (Dim1 × Dim2) factor map (e.g., participants Ha7 and Ha8 vis-à-vis the LE/Hb and NM groups, respectively; see Figure 10.2a) suggested that the additional socio-demographic variables in the MFA provided relevant data for understanding participant differences. Second, although participants within a group tended to resemble each other in terms of variable projections, even within a participant group, there were often considerable IDs for both demographic and phonetic profiles. Third, variables differed in terms of how big a role they played in distinguishing among participants. In the case of socio-demographic factors, for example, variables related to *what languages* are used between participants and their parents were more informative than variables related to *how often* they talk. In the case of phonetic factors, F_2 and VOT were more informative than F_1 or F_3, both in Mandarin and in English production,

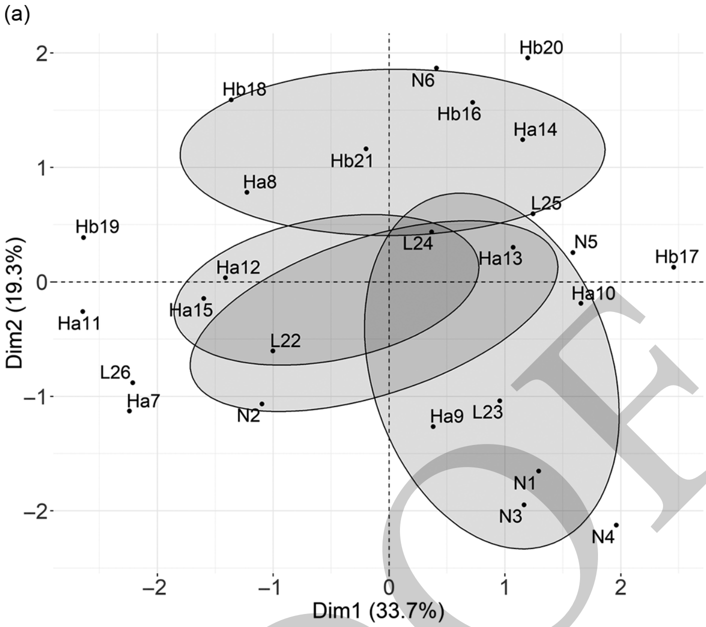

Figure 10.9 Representations of individuals according to English phonetic variables in MFA4. (a) Representations of individuals in terms of summed projections of all the variables in their English phonetic profile on Dim1 and Dim2 in MFA4 (ellipses mark the 95% CI around group means); (b) English phonetic profiles of four peripheral individuals (N3, L24, Ha15, Hb18) in terms of the five variable groups.

which is consistent with the focus on F_2 and VOT in prior work (Chang et al., 2011). These results suggest that for Mandarin–English bilinguals (in either order of acquisition), approximating target phonetic norms in their L2 may take longer for F_2 and VOT than for other phonetic properties, which could be due to the occurrence of problematic L1–L2 similarities in these areas that make it difficult for learners to form distinct L2 representations supporting target-like production (Flege, 1995).

Given the above three findings, we are optimistic about the potential for the MFA approach to open up avenues for integrating investigations of IDs in HLs with investigations of IDs in majority languages, which is rarely done, yet is necessary to gain a full picture of HSs as bilinguals. In the current study, carrying out side-by-side analyses of Mandarin and English production was facilitated by the common “currency” of the two analyses; each focused on the first two dimensions emerging from MFA, which incorporated a core of five variable groups for both languages. Such a design also allows for a limited

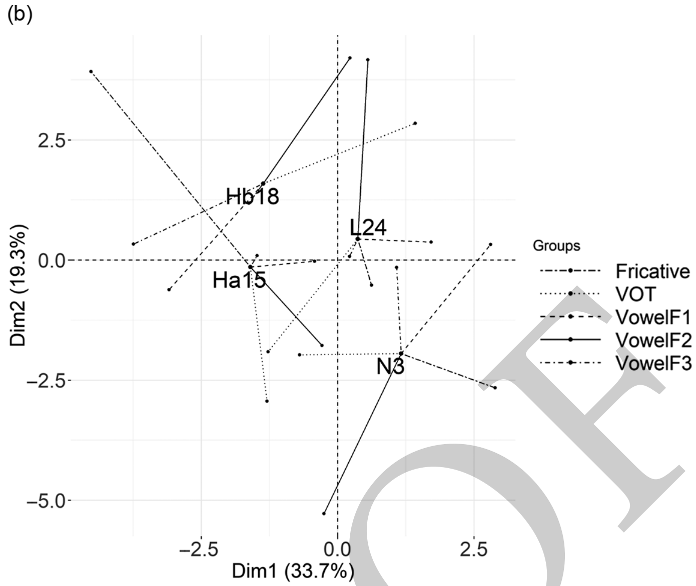

Figure 10.9 (*cont.*)

number of correlations that can be meaningfully compared across languages, which we present in this section as a demonstration of an approach to examining the relationship between demographic and phonetic profiles.

Focusing on the first two dimensions of each profile, our exploratory correlation analyses aimed to see whether dimensions of the demographic profile (DemoDim1, DemoDim2) can predict dimensions of the phonetic profile, both for Mandarin (ManDim1, ManDim2) and for English (EngDim1, EngDim2). These analyses showed that there was a significant, and strong, negative correlation of DemoDim1 with ManDim2 (Spearman's $\rho = -0.67$, $p < 0.05$), but not with ManDim1 (Spearman's $\rho = 0.29$, n.s.). All other correlations, including that of the demographic profile with the English phonetic profile, were not significant (Spearman's $\rho = -0.23$ to $\rho = 0.38$, n.s.).[6] In regard to the DemoDim1–ManDim2 correlation, recall that DemoDim1, starting from the low (negative) end, distinguished participants in the order L2, LE, HE, and NM, while ManDim2 was positively correlated with F_2 measures in back vowels and negatively correlated with VOT in aspirated stops. Thus, the negative DemoDim1–ManDim2 correlation reflects, in large part, the fact that, compared to the other groups, L2 participants produced higher F_2 values

[6] For correlation plots, see Figure 10 in the supplementary materials at https://osf.io/5hyjf/.

in /u ou/ and shorter VOTs for /p^h t^h k^h/, both of which place L2 participants at the high end of ManDim2. Considering the variable groups that projected onto DemoDim1, which included the trio *AoAr*, *SelfLg*, and *ParentsLg*, we interpret this result as supporting the view that both the timing and the manner of Mandarin exposure influence Mandarin production, particularly backness/ rounding in back vowels and long-lag VOT.

Although we are necessarily cautious about this apparent link between socio-demographic and phonetic dimensions for Mandarin but not English, as it is based on an MFA that considered a relatively small subset of socio-demographic variables that were available for all participants, this type of finding demonstrates that the outcomes of MFA can feed other types of analyses, leading to results that can be interpreted. Here, the correlation results provide suggestive evidence that socio-demographic variables are more critical for the HL than the majority language, consistent with a pattern in which HSs tend to be native-like more consistently in the majority language than in the HL (Chang, 2021).

In closing, we would like to synthesize what we have learned about group-based versus individual-centered approaches from the current study and our previous work with this same participant sample. Both approaches have advantages and disadvantages, so choosing between the two is a matter of evaluating trade-offs. On the one hand, the group-based approach is often easier to carry out, but risks obscuring important IDs. On the other hand, the individual-centered approach allows one to deeply analyze IDs, yet it may be prone to losing the “forest” for the “trees” (i.e., making larger generalizations harder to see). In our previous work, we took a group-based approach because we assumed distinct populations of Mandarin–English bilinguals, whose production patterns we wanted to distinguish empirically. The current study, however, has highlighted how constructs such as *native speaker*, *late learner*, and *heritage speaker*, even if readily understandable, can be difficult to distinguish from each other in practice; that is, these are categories with fuzzy boundaries, which means that a group-based study may end up drawing dividing lines that are, to some extent, arbitrary or artificial. By contrast, in the individual-centered approach, there is no need to draw preexisting dividing lines; instead, categories such as demographic groups can emerge from the data. This, we believe, is the chief advantage of the individual-centered approach; that is, it dispenses with dividing lines and, therefore, with the need to make decisions about group classifications.

This initial attempt at taking an individual-centered approach using MFA to shed light on variation in Mandarin HSs’ sound systems paves the way for several future research directions. For one, the current study was limited in its variables, which had been selected in previous work for a different purpose, and in its sample size of participants, which had not been determined with

MFA in mind. Thus, it would be useful in future research to expand the set of phonetic variables, for example, to represent the given sound system more comprehensively. Since a different, and larger, set of variables will probably change what projects onto each dimension in MFA, replicating this type of study with more variables and a larger sample size of participants will be essential to understanding the generalizability of the results. In addition, although we have exemplified analyses relating phonetic variables to socio-demographic variables in terms of the first two dimensions from MFA, there are other dimensions that can be explored, such as the third and fourth dimensions. These and other research avenues hold promise for providing a rich characterization of the remarkable diversity of HSs, both in their life experiences as well as in their sound systems.

Acknowledgments

We gratefully acknowledge helpful comments from the editor and three anonymous reviewers, as well as audiences at Western University, the University of Toronto, and the 10th International Symposium on the Acquisition of Second Language Speech (New Sounds 2022).

References

Abdi, H., Williams, L. J., & Valentin, D. (2013). Multiple factor analysis: Principal component analysis for multitable and multiblock data sets. *Wiley Interdisciplinary Reviews: Computational Statistics*, *5*(2), 149–179.

Boersma, P., & Weenink D. (2016). *Praat: Doing phonetics by computer* [Computer program]. Version 6.0.23. www.praat.org.

Bowles, A. R., Chang, C. B., & Karuzis, A. P. (2016). Pitch ability as an aptitude for tone learning. *Language Learning*, *66*(4), 774–808. https://doi.org/10.1111/lang.12159.

Chang, C. B. (2021). Phonetics and phonology of heritage languages. In S. Montrul & M. Polinsky (Eds.), *The Cambridge handbook of heritage languages and linguistics* (pp. 581–612). Cambridge University Press.

Chang, C. B., & Yao, Y. (2016). Toward an understanding of heritage prosody: Acoustic and perceptual properties of tone produced by heritage, native, and second language speakers of Mandarin. *Heritage Language Journal*, *13*(2), 134–160. https://doi.org/10.46538/hlj.13.2.4.

(2019). Production of neutral tone in Mandarin by heritage, native, and second language speakers. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Eds.), *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 2291–2295). Australasian Speech Science and Technology Association Inc.

Chang, C. B., Haynes, E. F., Yao, Y., & Rhodes, R. (2009). A tale of five fricatives: Consonantal contrast in heritage speakers of Mandarin. *University of Pennsylvania Working Papers in Linguistics*, *15*(1), 37–43.

(2010). The phonetic space of phonological categories in heritage speakers of Mandarin. In M. Bane, J. Bueno, T. Grano, A. Grotberg, & Y. McNabb (Eds.), *Proceedings from the 44th Annual Meeting of the Chicago Linguistic Society: The main session* (pp. 31–45). Chicago Linguistic Society.

Chang, C. B., Yao, Y., Haynes, E. F., & Rhodes, R. (2011). Production of phonetic and phonological contrast by heritage speakers of Mandarin. *Journal of the Acoustical Society of America*, *129*(6), 3964–3980. https://doi.org/10.1121/1.3569736.

Darcy, I., Park, H., & Yang, C.-L. (2015). Individual differences in L2 acquisition of English phonology: The relation between cognitive abilities and phonological processing. *Learning and Individual Differences*, *40*, 63–72. https://doi.org/10.1016/j.lindif.2015.04.005.

Debras, C. (2017). The shrug. *Gesture*, *16*(1), 1–34. https://doi.org/10.1075/gest.16.1.01deb.

Dewaele, J.-M., & Furnham, A. (2000). Personality and speech production: A pilot study of second language learners. *Personality and Individual Differences*, *28*(2), 355–365.

Dörnyei, Z., & Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Malden, MA: Blackwell.

Favier, S. (2020). *Individual differences in syntactic knowledge and processing: Exploring the role of literacy experience* [Doctoral dissertation, Max Planck Institute for Psycholinguistics, Netherlands].

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–272). York Press.

Hodge, G., Ferrara, L. N., & Anible, B. D. (2019). The semiotic diversity of doing reference in a deaf signed language. *Journal of Pragmatics*, *143*, 33–53. https://doi.org/10.1016/j.pragma.2019.01.025.

Idemaru, K., Holt, L. L., & Seltman, H. (2012). Individual differences in cue weights are stable across time: The case of Japanese stop lengths. *Journal of the Acoustical Society of America*, *132*(6), 3950–3964. https://doi.org/10.1121/1.4765076.

Iosad, P., & Lamb, W. (2020). Dialect variation in Scottish Gaelic nominal morphology: A quantitative study. *Glossa: A Journal of General Linguistics*, *5*(1), Art. 130. https://doi.org/10.5334/gjgl.1023.

Kartushina, N., & Frauenfelder, U. H. (2013). On the role of L1 speech production in L2 perception: Evidence from Spanish learners of French. In F. Bimbot, C. Cerisara, C. Fougeron, G. Gravier, L. Lamel, F. Pellegrino, & P. Perrie (Eds.), *Proceedings of the 14th Annual Conference of the International Speech Communication Association (INTERSPEECH-2013)* (pp. 2118–2122). ISCA Archive.

(2014). On the effects of L2 perception and of individual differences in L1 production on L2 pronunciation. *Frontiers in Psychology*, *5*, 1246. https://doi.org/10.3389/fpsyg.2014.01246.

Kassambara, A., & Mundt, F. (2020). *Package 'factoextra': Extract and visualize the results of multivariate data analyses (Version 1.0.7) [Computer software].* https://cran.r-project.org/package=factoextra.

Lê, S., Josse, J., & Husson, F. (2008). FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software*, *25*(1), 1–18. https://doi.org/10.18637/jss.v025.i01.

Milovanov, R., Pietilä, P., Tervaniemi, M., & Esquef, P. A. A. (2010). Foreign language pronunciation skills and musical aptitude: A study of Finnish adults with higher education. *Learning and Individual Differences*, *20*(1), 56–60.

Perrachione, T. K., Lee, J., Ha, L. Y. Y., & Wong, P. C. M. (2011). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *Journal of the Acoustical Society of America*, *130*(1), 461–472. https://doi.org/10.1121/1.3593366.

Piqueras-Fiszman, B., Ares, G., & Varela, P. (2011). Semiotics and perception: Do labels convey the same messages to older and younger consumers? *Journal of Sensory Studies*, *26*(3), 197–208. https://doi.org/10.1111/j.1745–459X.2011.00336.x.

R Development Core Team. (2021). *R: A language and environment for statistical computing* [Computer program]. In R Foundation for Statistical Computing. www.r-project.org/.

Raizada, R. D. S., Tsao, F.-M., Liu, H.-M., & Kuhl, P. K. (2010). Quantifying the adequacy of neural representations for a cross-language phonetic discrimination task: Prediction of individual differences. *Cerebral Cortex*, *20*(1), 1–12. https://doi.org/10.1093/cercor/bhp076.

Robinson, P. (2001). Individual differences, cognitive abilities, aptitude complexes and learning conditions in second language acquisition. *Second Language Research*, *17*(4), 368–392.

Schertz, J., Cho, T., Lotto, A., & Warner, N. (2015). Individual differences in phonetic cue use in production and perception of a non-native sound contrast. *Journal of Phonetics*, *52*, 183–204. https://doi.org/10.1016/j.wocn.2015.07.003.

Schertz, J., Cho, T., Lotto, A. J., & Warner, N. (2016). Individual differences in perceptual adaptability of foreign sound categories. *Attention, Perception, and Psychophysics*, *78*(1), 355–367. https://doi.org/10.3758/s13414-015-0987-1.

Skehan, P. (2002). Theorizing and updating aptitude. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 69–93). Amsterdam: John Benjamins.