Category and perceptual interference in second-language phoneme learning: An examination of English /w/-/v/ learning by Sinhala, German, and Dutch speakers

Paul Iverson[a*], Dulika Ekanayake[a], Silke Hamann[b], Anke Sennema[c]. and Bronwen G. Evans[a]

[a]Department of Phonetics and Linguistics, University College London, 4 Stephenson Way, London NW1 2HE, UK

[b]Utrecht University, Janskerkhof 13a, 3512 BL Utrecht, The Netherlands

[c]University of Potsdam, Institute of Linguistics, Karl-Liebknecht-Str. 24/25, 14476 Golm, Germany

[*]Corresponding author.

Abstract

The present study investigated the perception and production of English /w/ and /v/ by native speakers of Sinhala, German, and Dutch, with the aim of examining how their native language phonetic processing affected the acquisition of these phonemes. Subjects performed a battery of tests that assessed their identification accuracy for natural recordings, their degree of spoken accent, their relative use of place and manner cues, the assimilation of these phonemes into first-language categories, and their perceptual maps (i.e., multidimensional scaling solutions) for these phonemes. Most Sinhala speakers had near chance identification accuracy, Germans ranged from chance to 100% correct, and Dutch speakers had uniformly high accuracy. The results suggested that these learning differences were caused more by perceptual interference than by category assimilation; Sinhala and German speakers were similar in terms of first-language category assimilation, but the auditory sensitivities of Sinhala speakers made it harder for them to discern the acoustic cues that are critical to /w/-/v/ categorization.

Keywords: Category assimilation; perceptual interference; second language learning; speech perception; plasticity

Infants are born with language-universal perceptual abilities, being able to discriminate phonetic contrasts used in many languages (e.g., Eimas, Siqueland, Jusczyk, and Vigorito, 1971; Streeter, 1976). During the first year of life, their perceptual processes adapt to the acoustics of the ambient language, such that they become better at discriminating native-language (L1) phonetic contrasts (Kuhl, Stevens, Hayashi, Deguchi, Kiritani, and Iverson, 2006) and worse at discriminating some non-native phonetic contrasts (e.g., Werker and Tees, 1984). This perceptual specialization for L1 phonemes continues to develop into adolescence (Hazan and Barrett, 2000), and it has been theorized that this increasing specialization contributes to difficulties in second-language (L2) phoneme learning by adults (Flege, 2002; Kuhl, 2000; 2004).

The predominant view has been that this interference of L1 specialization on L2 perception and learning occurs at the level of segmental (i.e., phonological or phonetic) categorization. Individuals have long been thought to perceive L1 speech in terms of phoneme labels (e.g., Liberman, Harris, Hoffman, and Griffith, 1957), and to perceive L2 speech through the filter of their L1 phonological system (Trubetzkoy, 1969/1958). More recently, Best's Perceptual Assimilation Model (PAM; Best, 1994; Best, McRoberts, and Goodell, 2001) has suggested that listeners perceive L2 phonemes in terms of their articulatory similarity to L1 phonemes. For example, Japanese adults perceive the English /r/ and /l/ phonemes as both being weakly assimilated into their native Japanese /r/ category, and according to PAM this makes the English /r/ and /l/ phonemes sound the same (i.e., they both sound like poor examples of Japanese /r/; Best and Strange, 1992). In contrast, native English speakers are able to distinguish the Zulu voiceless-voiced distinction between the lateral fricatives /ɬ/ and /ɮ/ even though they do not have these

fricatives in their L1 (Best et al., 2001). English speakers are thought to be able to hear this difference because the phonemes are different in terms of how they are assimilated into L1 categories, with /ɬ/ sounding similar to voiceless English phonemes (e.g., /s/) and /ɮ/ sounding similar to voiced English phonemes (e.g., /l/).

Flege's Speech Learning Model (SLM; Flege, 1995; 2003) has suggested that these similarity relationships between L1 and L2 phonemes affect learning, because the L1 and L2 phonemes exist within the same phonological space, rather than being organized into independent subsystems. L2 categories are relatively easy to learn when they fall within an unoccupied region of the space (i.e., far from existing L1 categories). When a new L2 category is similar to an existing L1 category, individuals will often use the L1 category in their L2, and thus speak their L2 with a strong accent when the L1 category differs. Learning to reduce this accent often requires modification to one's L1 categories (e.g., creating merged or compromise categories that can accommodate both L1 and L2 phonemes), and this kind of change is easier when individuals use their L2 more extensively (Piske et al., 2001).

Although L1 segmental categorization almost certainly affects L2 phoneme learning to some extent, it is plausible that L1 exposure also alters auditory perceptual processing prior to the level of segmental categorization. For example, L1-specific patterns of perception have been found in the mismatch-negativity potential (MMN) of ERP and MEG experiments (Näätänen et al., 1997; Sharma and Dorman, 2000; Winkler et al., 1999), which is thought to reflect the resolution of pre-attentive auditory processing (e.g., Näätänen, 2001). Moreover, the developmental evidence suggests that infants begin to exhibit language-specific patterns of perceptual sensitivity around 6 months of age

(Cheour et al., 1998; Kuhl and Coffey-Corina, 2001; Kuhl et al., 1992), prior to the age at which they develop phonological or lexical categories. Kuhl (2000; 2004) has suggested that the perceptual abilities of infants become tuned to the statistical distribution of the ambient speech that they hear, such that they lose perceptual sensitivity near distribution peaks (i.e., prototypes; Iverson and Kuhl, 1995, 1996, 2000; Kuhl, 1991), and that these perceptual changes facilitate later category learning.

Our work (Iverson, Kuhl, Akahane-Yamada, Diesch, Tohkura, Kettermann, and Siebert, 2003; Iverson, Hazan, and Bannister, 2005) suggests that these sorts of distortions in auditory-phonetic processing may also affect L2 learning by adults, in ways that cannot be easily explained by segmental categorization. For example, Japanese adults have a marked difficulty learning the English /r/-/l/ category (Goto, 1971; Miyawaki et al, 1975); learners can benefit from experience and training (e.g., Logan et al., 1991; Lively et al., 1993) but it can take decades of experience to achieve native-like perception and production (Flege, Takagi, and Mann, 1995). From the standpoint of SLM (Flege, 1995, 2003), this level of difficulty is puzzling because the closest Japanese phoneme (a lateral flap /r/) is only moderately related to English /r/ and /l/. To help explain this difficulty, we mapped Japanese adults' perceptual sensitivities for these phonemes in a two-dimensional stimulus space composed of second (F2) and third formant (F3) frequencies. Contrary to Best and Strange (1992), we found that Japanese adults were quite able to hear acoustic variation among English /r/ and /l/ stimuli, but they were more sensitive to acoustic dimensions that were irrelevant to the categorization of English /r/ and /l/ (F2) than they were to the acoustic variation that is critical to categorization by native speakers (F3 differences at the /r/-/l/ boundary). We concluded

that Japanese adults have an underlying perceptual space for /r/ and /l/ that *perceptually interferes* with L2 category learning by making the critical acoustic variation less salient. In contrast, native English speakers have perceptual spaces for these stimuli that facilitate categorization, with low sensitivity near the best exemplars of /r/ and /l/, and high sensitivity at the category boundary (Iverson and Kuhl, 1996).

The present study extended our investigation to the perception of the English /w/-/v/ distinction by L1 speakers of Sinhala, German, and Dutch. Sinhala and German both have one phoneme that is similar to English /w/ and /v/, and Dutch has two related phonemes. Sinhala has /ʊ/, a labiodental approximant[1] that has the manner of English /w/ but the place of English /v/. German has a voiced labiodental /v/; the German /v/ is usually described as a fricative as is English /v/ (Kohler, 1999), but German speakers often produce little contact for labiodental sounds, particularly in initial position, and thus can have an approximant-like realization of the sound that makes it similar to /ʊ/ (see Scherer and Wollmann, 1985). Dutch speakers, depending somewhat on their dialect, have a labiodental approximant /ʊ/, and a labiodental fricative /v/ that is sometimes devoiced to be realized as similar to /f/ (Booij, 1995; Gussenhoven, 1999; Hamann and Sennema, 2005).

The aim of this study was to test whether L1 category assimilation and perceptual interference can account for the degree of difficulty that these language groups have in learning these phonemes. Informal reports from native speakers indicated to us that Sinhala and German speakers differ in their ease of learning English /w/-/v/, despite the fact that both of these groups have one related L1 phoneme; Sinhala speakers can find it very difficult to produce this distinction even after decades of living in an English

speaking country, but many German speakers seem to be able to learn this contrast through classroom instruction. Dutch speakers have little reported difficulty hearing the English /w/-/v/ distinction, despite the fact that their native phonemes are not an exact match for English /w/ and /v/. Our subject selection was designed with the aim of finding a range of good and poor English /w/-/v/ identification scores within each language group. That is, we tested Sinhala speakers who were highly experienced with English (i.e., long-term residents of London), because we wanted to find at least a subset of subjects who were accurate at /w/-/v/ identification. In contrast, we tested German speakers with less English experience (i.e., German university students, few of whom had lived abroad) because we wanted to find at least a subset of subjects who had difficulty with /w/-/v/. We also tested Dutch speakers with less experience, but were unable to find subjects with /w/-/v/ identification difficulty.

In order to verify how well L1 speakers of these languages were able to learn the English /w/-/v/ contrast, the baseline abilities of subjects were assessed by having subjects identify English /w/ and /v/ for natural recordings from multiple talkers, and produce this English contrast when reading an accent-revealing sentence. Their categorization of these consonants was assessed by presenting a synthetic /ɑ:wɑ:/-/ɑ:vɑ:/ grid of stimuli that varied orthogonally in place and manner, and having individuals identify and rate the goodness of these stimuli both in terms of their L1 and L2 (English) phonemes. Their underlying perceptual space for these phonemes was examined via multidimensional scaling (MDS) of similarity ratings for stimulus pairs from the synthetic grid, in combination with same-different discrimination judgements for selected pairs. The results were compared to determine the extent to which L1

assimilation patterns and the underlying perceptual space can account for the identification accuracy of natural stimuli.

Method

Subjects

The subjects were 20 native Sinhala speakers, 22 native German speakers, and 18 native Dutch speakers, as well as 20 native English speakers who were tested to provide normative data. Two subjects (1 Sinhalese, 1 German) were dropped from the data analysis because their identification of natural stimuli was markedly below chance (< 35% correct), which indicated that they may have heard a difference between /w/ and /v/ but reversed the response labels.

The Sinhala speakers were tested in London; they were 21-71 years of age (median = 54), began learning English when they were 2.5-18 years of age (median = 5), moved to England when they were 14-40 years of age (median = 24), and had lived in London for 0.5-41 years (median = 28). The German speakers were tested in Potsdam (Germany); they were 19-46 years of age (median = 23) began learning English when they were 6-44 years of age (median = 11), and none had lived in an English speaking country for more than 1 year. The Dutch speakers were tested in Nijmegen (The Netherlands) and Berlin; they were 17-64 years of age (median = 20), began learning English when they were 8-14 years of age (median = 11), and none had lived in an English speaking country for more than 1.6 years. The native English speakers were tested in London, and were 20-60 years of age (median = 28). All subjects reported having no known hearing impairments or learning disabilities.

Stimuli and apparatus

Each subject was tested in a quiet room with stimuli delivered over headphones; responses were recorded via a graphical interface presented on either a PC or a Pocket PC. Digital recordings were made of each subject's voice, with a minimum of 21,050 16-bit samples per second.

A test corpus of native-language /w/ and /v/ recordings were made from 4 talkers (2 male and 2 female). The consonants were embedded in VCV syllables, with vowel contexts of /i:/-/i:/, /ɑ:/-/ɑ:/, and /u:/-/u:/. The stimuli were designed to have relatively high variability (i.e., multiple talkers and vowel contexts) to increase the likelihood that the results would generalize to other talkers and contexts, and to increase the difficulty of the identification task. There were a total of 48 stimuli (2 consonants X 4 talkers X 3 syllables X 2 recordings of each). All recordings were made in an anechoic chamber with 44,100 16-bit samples per second.

A set of 16 synthetic /ɑ:/-/ɑ:/ stimuli varied 2 dimensions (manner and place), with 4 steps along each dimension (see Figure 1). The synthesis parameters were initially set to model a natural recording of /ɑ:vɑ:/ from a female speaker. Pilot testing was used to determine which acoustic parameters to vary during the consonant in order to create realistic versions of /w/-/v/. The place distinction varied F2 along 4 steps that were equally spaced on the ERB scale (680, 875, 1112, and 1400 Hz, from /w/ to /v/; Moore, Glasberg, and Baer, 1997). The manner distinction could not be made using a single acoustic dimension (native English pilot subjects varied in their use of acoustic cues), so F1, amplitude of frication, and transition duration were covaried along this stimulus dimension. F1 was varied from high to low for /w/-v/, with equally spaced steps along the

ERB scale (316, 267, 223, 182 Hz). Amplitude of frication was varied from low to high for /w/-/v/ (-34, -31, -28, -25 dB, relative to the amplitude of voicing during the closure). The duration of the transition to the following vowel was varied from long to short for /w/-/v/ (80, 67, 53, 40 ms) and the duration of the closure was varied inversely (60, 73, 87, 100 ms) so that the closure and following transition duration always had a total length of 140 ms. All other parameters (e.g., the amplitude of voicing, and the vowel parameters) were the same for all stimuli.

Procedure

Natural stimulus identification. On each trial, subjects heard a naturally recorded VCV syllable and identified whether the consonant was /w/ or /v/. They did not receive feedback and were not able to replay the stimulus. Subjects first completed a practice block of 16 trials, then completed an experimental block of 48 trials (each stimulus presented once, in a random order). This task was completed by native Sinhala, German, and Dutch speakers, but not native English speakers.

Voice recordings. All subjects were recorded reading the accent-revealing sentence, "The heavy wind swept away Valerie's velvet scarf." Subjects were able to familiarize themselves with the sentence (i.e., read the sentence without speaking), but did not rehearse the sentence before making their recording. After all subjects were tested, 4 phonetically trained native English speakers rated the recordings for accent. They gave three numerical ratings (1-7) for each sentence: the degree of non-native accent for /w/, the degree of non-native accent for /v/, and the degree of contrast between /w/ and /v/. For example, some Dutch speakers produced their English /w/ without velarization and their English /v/ with little voicing (i.e., like /f/); such speakers were rated as producing

each consonant with a strong accent but they were rated as having a high degree of contrast (i.e., the consonants sounded different, even though neither were native-like). Other speakers produced a /v/ phoneme for both English /w/ and /v/; such speakers were rated as having a strong accent for /w/, little accent for /v/, and no contrast between /w/ and /v/.

English identification and goodness ratings. On each trial, subjects heard one of the synthetic syllables, identified whether the consonant was English /w/ or /v/, and then rated on a continuous graphical scale whether the stimulus was a good or poor exemplar of the phoneme that they identified. That is, if they identified the stimulus as /v/, they then rated whether the stimulus was a good or poor exemplar of the English /v/ phoneme. Subjects heard the stimulus once before making their identification judgment, but were able to repeat the stimulus as many times as they wanted before making their goodness judgment. Subjects completed a practice block of 16 trials (each stimulus played once in a random order), and then an experimental block of 64 trials (each stimulus played 4 times, in a random order).

L1 identification and goodness ratings. This task was similar to the English identification and goodness rating task, except that each group of listeners identified and rated the goodness of the synthetic stimuli in terms of their own native-language phonemes (native English speakers did not participate). Subjects were told that the stimuli were synthesized based on English, but they needed to judge how the stimuli would be interpreted if they heard it spoken in their native language (e.g., if they were listening to a native English speaker who was learning to speak Dutch). Each group had two identification choices; Sinhala speakers identified these stimuli as their /ʊ/ or as "out

of category," German speakers identified these stimuli as their /v/ or as "out of category," Dutch speakers identified these stimuli as their /ʊ/ or as their /v/. Subjects gave a goodness rating for stimuli that they thought were members of a native category, but did not give a goodness rating after "out of category" responses. Subjects completed a practice block of 16 trials (each stimulus played once in a random order), and then an experimental block of 64 trials (each stimulus played 4 times, in a random order).

Similarity scaling. Subjects heard two stimuli with a 250 ms ISI on each trial, and rated the similarity of the stimuli on a continuous graphical scale from same to different. The stimulus corpus comprised all possible pairs of the 16 synthesized stimuli, with no stimulus paired with itself, and each pair presented in both orders (240 pairs). Subjects started with a practice of 16 randomly selected pairs. Subjects then completed an experimental block comprising all 240 pairs presented in a random order.

Discrimination. Subjects heard two stimuli with a 250 ms ISI on each trial, and judged whether the stimuli were same or different. The stimuli were the 4 items along the diagonal from /w/ to /v/ in the synthesized set of 16 stimuli (i.e., lower-left corner to upper-right corner in Figure 1, with place and manner covaried). Same trials had the same stimulus presented twice. Different trials had two neighboring stimuli along the diagonal; the 4-step continua thus had 3 different pairs. Subjects completed a practice block comprising 4 same trials (one for each stimulus) and 6 different trials (each of the pairs in both orders), with the order of trials randomized. Subjects then completed an experimental block comprising 20 same trials and 30 different trials in a random order (i.e., 5 repetitions of each pair in the practice).

Results and Discussion

Natural stimulus identification

As displayed in Figure 2, there were substantial differences between the groups in terms of their identification accuracy for natural speech. The majority of Sinhalese subjects were near chance at identifying English /w/ and /v/, with only three subjects performing higher than 75% correct. The majority of German subjects were accurate at identifying English /w/ and /v/ (i.e., median accuracy = 94%), but the lower quartile of subjects had poor accuracy, ranging from 46-77% correct. All Dutch speakers were more than 90% accurate at identifying these consonants. A one-way ANOVA confirmed that the differences between language groups were significant, $F(2,55) = 47.20$, $p < 0.001$.

The low accuracy of the Sinhalese speakers was particularly striking given that these subjects had extensive experience with English. The benchmark for difficult L2 phoneme learning has been Japanese adults learning the English /r/-/l/ distinction, and even this contrast can be learned to native-like levels after extensive exposure (e.g., living in an English-speaking country for 20 years; Flege et al. 1995). The present results demonstrate that most Sinhala speakers are near chance at English /w/-/v/ even after having a median of 28 years of experience living in an English-speaking country, and having English lessons beginning at a median of 5 years of age. This case is thus a rather extreme example of L2 phoneme learning difficulty.

An inspection of the Sinhalese responses revealed that they were heavily influenced by the vowel context. When the context was /i:/, subjects were 79% correct for /v/ stimuli and 39% correct for /w/ stimuli. When the context was /u:/, subjects were 39% correct for /v/ and 79% correct for /w/. When the context was /ɑ:/, subjects were 48% correct for /v/

and 66% correct for /w/. Their response biases thus changed with the second formant (F2) of the following vowel; listeners were biased to identify stimuli as /v/ when F2 was high (/i:/), and biased to identify stimuli as /w/ when F2 was low (/u:/ and /ɑ:/).

For graphical purposes in later Figures (but not for inferential statistics), the subjects were split based on their identification accuracy for natural stimuli (greater or less than 75%). This 75% criterion was chosen because it was halfway between chance and 100%, it separated the 3 outlier high-accuracy Sinhala speakers from the rest, and it separated the lowest-quartile of German speakers from the rest. The data from Dutch speakers was not split in this way because all had identification accuracy well above this criterion. The data from English speakers was also not split because the stimuli had been screened to be clearly intelligible by native speakers.

Accent ratings

The accent ratings for the Sinhala, German, and Dutch speakers are displayed in Figure 2. English speakers uniformly received high accent ratings (median of 7 on all measures), so these scores are not displayed. When speaking *The heavy wind swept away Valerie's velvet scarf*, Sinhala speakers were judged to produce a /v/-like phoneme for both English /w/ and /v/; they were rated to have produced a low degree of contrast between /w/ and /v/, their /w/ production was rated to have a strong non-native accent, and their /v/ production was judged to have a milder non-native accent. German and Dutch speakers had only slight non-native accents for /w/; both groups produced these phonemes with somewhat less velarization or lip-rounding compared to native speakers. German and Dutch speakers also had similar degrees of non-native accent for /v/; Dutch speakers tended to have more frication and less voicing (i.e., more like English /f/) than

did native English speakers, and Germans tended to produce /v/ with less frication (i.e., more like /w/) than did native English speakers. Thus, Germans and Dutch speakers had very similar degrees of non-native accent, but Dutch speakers produced English /w/ and /v/ more contrastively.

One-way ANOVAs demonstrated that the language groups were significantly different in terms of degree of contrast, $F(2,55) = 37.78$, $p < 0.001$, and /w/ accent, $F(2,55) = 49.13$, $p < 0.001$; the groups did not differ in terms of /v/ accent, $p > 0.05$. Across all L2 speakers, the individual differences in identification accuracy were significantly correlated with the degree of contrast, $r = 0.74$, $p < 0.001$, and the degree of accent for /w/, $r = 0.71$, $p < 0.001$. However, none of these correlations were significant, $p > 0.05$, when calculated separately for each language group, and the accent ratings for /v/ were not significantly correlated with identification when calculated either across or within these groups. There was thus some indication of a link between spoken accent and identification accuracy, but this link was largely driven by language group differences (i.e., Dutch speakers, as a group, having low accent and high identification accuracy, and Sinhala speakers having the reverse) rather than by finer-grained individual differences.

English identification and goodness ratings

As displayed in Figure 3, native English speakers used both place and manner to distinguish /w/ and /v/; there was a diagonal category boundary, and the best exemplars of the /w/ and /v/ categories were in the upper-right and lower-left corners of the stimulus space. The L2 speakers who had relatively high recognition accuracy for natural stimuli (>75%), gave identification and goodness ratings for the synthetic stimuli that were relatively similar to those of native English speakers, with perhaps some fine-grained

differences in how these listeners used these acoustic cues. For example, Dutch speakers had a somewhat more horizontal identification boundary, suggesting that they based their judgments on manner more than did native English speakers. Individuals who were relatively inaccurate at identifying natural speech (<75% correct responses) did not have a clear identification boundary for these stimuli. For example, the /w/ identification percentages for the less-accurate Sinhala speakers ranged from 47 to 66% for individual stimuli, and the goodness judgments ranged from 0.48 to 0.66 on a scale from 0 (bad) to 1 (good); none of the stimuli were consistently categorized as /w/ or /v/. The pattern for the less-accurate German speakers was similar, except that there was a tendency to give lower goodness ratings to stimuli in the lower-right corner of the stimulus space.

To compare the English identification and goodness judgments, category centroids were calculated for each subject for /w/ and /v/. Each centroid measured the location of the "center" of each subjects /w/ or /v/ judgments within this stimulus space, and was calculated by weighting the position of each stimulus by its identification and goodness judgments. For example, the centroid for /w/ responses along the manner dimension was calculated using the formula

$$/w/-centroid_{manner} = \frac{\sum_{i=1}^{4}\sum_{j=1}^{4} i p_{ij} g_{ij}}{\sum_{i=1}^{4}\sum_{j=1}^{4} p_{ij} g_{ij}}, \quad (1)$$

where i is the manner step number, j is the place step number, p is the probability that this stimulus was identified as /w/, and g is the average /w/ goodness rating for that stimulus. Analogous formulas were used for /v/ and for place.

The distance between each individual's /w/ and /v/ category centroids was used as an index of how consistently they divided these stimuli into two categories. The /w/ and /v/ centroids were close together and in the middle of the stimulus space for subjects whose distribution of /w/ and /v/ responses overlapped (e.g., Sinhala speakers with natural identification < 75%), but were far apart for subjects who had a clear boundary between /w/ and /v/ (e.g., Dutch speakers). A one-way ANOVA demonstrated that this centroid distance was significantly different between all four language groups, $F(3,74) = 36.26$, $p < 0.001$, with Sinhala speakers having substantially less separation between their centroids. Across the L2 language groups, the distance between category centroids was significantly correlated, $r = 0.86$, $p < 0.001$, with the identification percentages for natural stimuli. Within L2 language groups, the two measures were significantly correlated for Sinhala, $r = 0.85$, $p < 0.001$, German, $r = 0.68$, $p < 0.001$, and Dutch speakers, $r = 0.47$, $p = 0.049$. The results thus demonstrate that their judgments on these synthetic stimuli were consistently related to their identification of natural speech.

To examine the relative weightings of place and manner cues, the distance between the /w/ and /v/ centroids along the place dimension was subtracted from the distance between centroids along the manner dimension. A one-way ANOVA demonstrated that the four language groups differed significantly on this measure, $F(3,74) = 6.16$, $p < 0.001$; English, Sinhala, and German speakers gave nearly equal weight to place and manner, but Dutch speakers put significantly more weight on the manner dimension. However, this relative weighting measure was not significantly correlated, $p > 0.05$, with natural identification accuracy, either across or within language groups.

L1 identification and goodness ratings

One of the central claims of SLM (Flege, 1995, 2003) is that individuals use the same category representations for L1 and L2 phonemes, rather than learning independent phonological systems. This hypothesis was mostly confirmed by the L1 identification and goodness ratings (see Figure 4). For example, Dutch speakers used a diagonal identification boundary when categorizing these stimuli as Dutch /ʊ/ and /v/ that matched their boundary when categorizing these stimuli as English /w/ and /v/. Paired t-tests were used to compare the category centroids for their L1 and L2 categories. There was a marginally significant difference between Dutch /ʊ/ and English /w/ along the manner dimension, $t(17) = -2.06$, $p = 0.055$, suggesting that their L1 /ʊ/ may have had a category centroid that was slightly lower on the manner dimension compared to their L2 /w/ (mean difference of 0.11 steps). However, there were no significant differences in the /w/ centroids along the place dimension, and no differences between the /v/ centroids along either dimension, $p > 0.05$.

The more-accurate German speakers (i.e., identification > 75%) also gave parallel responses for their L1 and L2; they gave similar responses for their L1 and L2 /v/, and judged that the stimuli that sounded like English /w/ were not examples of any German phoneme. It was less clear whether the less-accurate German speakers responded to the stimuli the same way in their L1 and L2; they consistently judged that stimuli toward the upper-middle of the stimulus space were good exemplars of German /v/, but these same stimuli were not consistently identified as English /v/. However, paired t-tests across all German speakers revealed no significant differences between the L1 and L2 /v/ category centroids along either the place or manner dimension, $p > 0.05$.

The less-accurate Sinhala speakers judged that these stimuli were all mediocre members of their L1 /ʊ/ category, with averages of 59-83% /ʊ/ identification and 0.53-0.65 goodness; this parallels their inconsistent identification of these stimuli as English /w/ and /v/. There were no significant differences between their Sinhala /ʊ/ and English /v/ centroids, $p > 0.05$, but their Sinhala /ʊ/ was significantly different from their English /w/ in both manner, $t(20)=8.38$, $p < 0.001$, and place, $t(20)=6.15$, $p<0.001$. This suggests that Sinhala speakers, as a group, responded to English /v/ much in the same way as they did for their L1 /ʊ/ category. However, the 3 Sinhala speakers with identification accuracy > 75% appeared to have a different pattern of responses. Their L1 /ʊ/ identifications were closer to their English /w/ identifications, with a trend to more consistently identify English /w/ when the place was labiovelar and Sinhala /ʊ/ when the place was labiodental. Stimuli that were identified as English /v/ were judged to be on the fringes of the Sinhala /ʊ/ category. Although we were not able to find enough more-accurate Sinhala speakers to test this statistically, the result suggests that these subjects may have different patterns of assimilation than the less-accurate Sinhala speakers.

Multidimensional scaling and discrimination

Kruskal (1964) non-metric MDS was used to analyze the similarity ratings and fit the items into 2-dimensional spaces (see Figure 5). The stress values had a range of 14.4-5.62, accounting for 86-97% of the variance in the original similarity ratings. The two groups of Sinhala speakers had somewhat weaker fits to their data; all other solutions accounted for greater than 94% of the variance. MDS solutions typically display only the relative similarity within a stimulus set, so the discrimination results were used to scale the solutions to make them reflect the absolute similarity. Specifically, the cumulative *d'*

(a perceptual distance measure that can be calculated from discrimination data using Detection Theory; Macmillan and Creelman, 1991) was calculated for the diagonal from /w/ to /v/, and the sizes of the MDS solutions were scaled to reflect these differences. For example, the cumulative *d'* for English subjects (mean = 6.84) was 2.3 times greater than for the low-accuracy Sinhalese subjects (mean = 2.29), so the MDS solutions were scaled such that the distance between the ends of that diagonal was 2.3 times larger for English subjects than for the low-accuracy Sinhalese subjects.

As displayed in Figure 5, the listeners varied in their overall ability to distinguish these stimuli. Although the low-accuracy Sinhalese subjects were able to discriminate these stimuli above chance, their discrimination ability was much lower than for the English, Dutch, or high-accuracy German subjects. A one-way ANOVA confirmed that cumulative *d'* was significantly different between language groups, $F(3,74) = 12.23$, $p < 0.001$. The cumulative *d'* was significantly correlated with individual differences in the identification of natural stimuli across L2 language groups, $r = 0.60$, $p < 0.001$. However, none of these correlations were significant within language groups, $p > 0.05$.

The MDS solutions also differed in terms of their stretching in the middle of the perceptual space (i.e., stimuli that straddled the native English /w/-/v/ boundary). English, Dutch, and high-accuracy German subjects all had relatively poor discrimination sensitivity within the native English /w/ and /v/ categories but had relatively high sensitivity in both manner and place at the category boundary. High-accuracy Sinhalese subjects had a stretched perceptual space at the category boundary, but only along the manner dimension. Low-accuracy Sinhalese subjects had no evidence of stretching at the category boundary, and low-accuracy German subjects appeared to have somewhat less

stretching at the category boundary than did high-accuracy subjects. The degree of stretching was quantified for individual subjects by calculating *d'* for the middle pair of the diagonal (i.e., the pair that crossed the native English /w/-/v/ boundary) and subtracting the *d'* from the upper-right pair of the diagonal (i.e., stimuli that were clearly within the native English /v/ category). A one-way ANOVA confirmed that the degree of stretching was significantly different between language groups, $F(3,74) = 10.94$, $p < 0.001$. The degree of stretching was significantly correlated with identification accuracy for natural speech across language groups, $r = 0.44$, $p < 0.001$. Within language groups, the correlation was only significant for Sinhalese subjects, $r = 0.56$, $p = 0.013$.

One of the key claims of PAM (Best, 1994; Best et al., 2001) is that speech is perceived in terms of phonological categories, such that two stimuli will sound the same if they are categorized as being equally good exemplars of the same L1 phonological category. This prediction is consistent with the Sinhalese data. Specifically, the low-accuracy Sinhalese subjects heard these stimuli as all being equally mediocre exemplars of the Sinhala /ʊ/, so PAM would predict that these stimuli should all sound the same. The MDS and discrimination results come close to showing this pattern; subjects did not literally hear no difference between the stimuli (i.e., *d'* was > 0), but their cumulative sensitivity was half that of native English speakers. The high-accuracy Sinhalese subjects perceived that L1 category assimilation varied along the manner dimension, with goodness and /ʊ/-categorization declining toward the fricative end of the dimension. Their MDS solutions were likewise stretched along the manner dimension as the stimuli moved out of the L1 category, suggesting that they may have perceived these stimuli in terms of L1 goodness.

However, category assimilation cannot explain the MDS solutions of the other listener groups. For example, the goodness and identification of stimuli for Dutch speakers varied along the lower-left to upper-right diagonal (i.e., from /w/ to /v/), but the orthogonal diagonal (upper left to lower right) had little effect on goodness and identification. If subjects had been perceiving these stimuli in terms of category goodness, the space should have collapsed the upper-left to lower-right diagonal because these stimuli all sound the same with regard to category membership. In contrast, the obtained MDS solution was two-dimensional; subjects were able to discern variation along the orthogonal diagonal that had little effect on categorization or goodness. The high-accuracy German and English spaces are problematic for PAM (Best, 1994; Best et al., 2001) in the same way. Listeners are clearly able to hear acoustic variation along more dimensions than their one-dimensional /w/-/v/ categorization.

The predictions of category assimilation are even less consistent with the MDS solution of the low-accuracy German subjects. These subjects heard the upper-center of the stimulus space as having good exemplars of the German /v/ category, with goodness and identification declining in both directions toward the edges of the stimulus space. PAM (Best, 1994; Best et al., 2001) would predict a complex distortion of the perceptual space, with the left edge of the space collapsing onto the right edge of the space, because both sound the same in terms of their German /v/ categorization. However, there is no evidence in the MDS solutions that these two edges of the space sounded the same.

General Discussion

The results confirmed that L1 speakers of Sinhala, German, and Dutch differed in their ability to perceive and produce the English /w/-/v/ distinction. Despite the fact that

the Sinhala speakers were very experienced with English (started to learn English at a median of 5 years old, moved to England at a median of 24 years old, and had lived in England for a median of 28 years), most of these individuals were near chance at identifying natural English /w/ and /v/, spoke these consonants with a poor degree of contrast, and were poor at discriminating acoustic differences between these phonemes. There was high variability among German subjects, but on average they identified /w/-/v/ accurately, spoke these phonemes with a high degree of contrast, and discriminated these stimuli as well as native English speakers. Dutch speakers were more uniformly accurate in their identification accuracy, production contrast, and discrimination.

Why did these three language groups differ in their perception and production of English /w/-/v/? Some of the ability of Dutch speakers to accurately identify English /w/ and /v/ may have been due to their experience with English (e.g., English television programs are not dubbed into Dutch) or motivation to learn new languages, but it is notable that their degree of spoken accent was as strong as for German speakers. Dutch speakers assimilated English /w/ into their L1 Dutch /ʊ/ category and English /v/ into their L1 Dutch /v/ category. The difference between Dutch /ʊ/ and /v/ is primarily a manner distinction, and they likewise gave more weight to the manner dimension, compared to the other language groups, when categorizing English /w/-/v/. The evidence thus suggests that these speakers simply used their existing L1 Dutch phonetic categories when perceiving and producing English, in accord with the predictions of SLM (Flege, 1995; 2003). Moreover, their underlying perceptual space for these phonemes likely facilitated rather than interfered with the English categorization; Dutch speakers had relatively high sensitivity near the English /w/-/v/ boundary. The fact that they could hear

differences between English /w/ and /v/ was also broadly in accord with PAM (Best, 1994; Best et al., 2001), because these stimuli mapped onto two different L1 categories. Thus, both category assimilation and perceptual interference seemed to promote their categorization of English /w/ and /v/.

However, categorization models cannot easily explain why German speakers were better than Sinhala speakers at learning the English /w/-/v/ distinction. Both groups of speakers had one L1 phoneme that was similar to English /w/ and /v/, and thus they needed to acquire at least one new category in order to distinguish these phonemes. SLM (Flege, 1995, 2003) predicts that acquiring a new category should be easier when it is far away from any existing category, but German /v/ actually appeared to be somewhat closer to English /w/ and /v/, in that there was a good German /v/ in this stimulus space but none of the stimuli were consistently identified as a very good Sinhala /ʊ/.

Several theories (e.g., Flege, 1995, 2003; Kuhl, 2000, 2004) predict that new L2 phonemes will be learned better when the age of learning is younger and when individuals have more experience with their L2. On the surface, the present results also conflict with expectations, because the Sinhala speakers started earlier and had far more experience than did the Germans. However, it is worth noting that the Sinhala speakers did not move to England until they were a median of 24 years old; it is plausible that they had been mostly exposed to an English accent in Sri Lanka that had not distinguished /w/ and /v/, and they may thus have had relatively late exposure to speakers who used /w/ and /v/ contrastively. That being said, their median 28 years of living in England would have been expected to compensate for this late exposure; Japanese speakers with similar

histories and analogous patterns of L1 phonological assimilation are able to learn to produce English /r/ and /l/ accurately (Flege et al., 1995).

Perceptual interference (Iverson et al., 2003) can more directly explain these differences between Sinhala and German speakers. The Sinhala speakers with poor identification accuracy had a perceptual space for these stimuli that interfered with learning; they had poor overall levels of sensitivity to acoustic differences as well as having no local region of high sensitivity at the /w/-/v/ category boundary. Such a perceptual space would be hard to alter because it would be self-reinforcing. That is, even if they were trained on clear English /w/-/v/ stimuli, their perceptual space would make the stimuli sound the same and thus promote listeners to continue to process them the same. In contrast, the perceptual space of the German speakers with poor identification accuracy would not greatly interfere with learning; they are able to hear the difference between English /w/ and /v/ more similarly to L1 speakers, albeit with somewhat less stretching at the category boundary. From the standpoint of perceptual interference, the less-accurate Germans should thus be capable of learning the English /w/-/v/ distinction; their poor identification accuracy may have reflected their educational experience or motivation to learn rather than an inherent lack of plasticity.

Why did the Sinhala speakers have such poor perceptual sensitivity? Kuhl's Native Language Neural Commitment model (e.g., Kuhl, 2000, 2004; c.f. Werker and Yeung, 2005) suggests that the perceptual sensitivities of infants adapt to the statistical distributions of ambient speech during the first year of life, such that they lose sensitivity near distribution peaks (i.e., prototypes) and gain sensitivity in the troughs between the distributions. These perceptual sensitivities are thought to facilitate early stages of word

and phonological learning, and the acquisition of phonological categories can in turn have effects on perceptual sensitivities, such that perceptual sensitivity and categorization become coupled. We do not have a means of comparing the statistical distribution of phonemes heard by Sinhalese and German infants. However, it is notable that the Sinhala speakers began learning English at a relatively young age (median 5 years old). If they were taught a variety of English that did not contrast /w/ and /v/ (e.g., if they primarily heard the speech of Sinhala-English bilinguals that did not make this distinction), then this would promote them to lose perceptual sensitivity to the difference between these phonemes. They may have thus altered their perceptual processing at a young age in a way that blocked learning when they were exposed as an adult to English accents that had a /w/-/v/ contrast.

Although developmental processes such as these may tend to couple categorization and perceptual sensitivity, this does not mean that stimuli are literally perceived in terms of phonological categories. For example, categorization models, such as PAM (Best, 1994; Best et al., 2001), can account for perceptual differences along single vectors between phoneme categories (e.g., along a continuum from good /w/ to good /v/), but it seems clear that individuals can perceive acoustic variation along multiple phonetic dimensions and that distortions along these dimensions affect learning (see Iverson et al., 2003). The present results do not invalidate the basic idea that L1 phonological categories have an important role in L2 phonological learning and perception, but they suggest that distortions in the underlying perceptual space can additionally contribute to the L2 learning process.

Acknowlegements

We gratefully acknowledge funding by the UK Economic and Social Research Council (RES-000-23-0838) and the German Science Foundation (DFG; GWZ 4/8-1-P2 for Silke Hamann and SFB 632-C4 for Anke Sennema).

References

Best, C. T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J. C. Goodman & H. C. Nusbaum (Eds.), The development of speech perception: The transition from speech sounds to spoken words (pp. 167-224). Cambridge, MA: MIT Press.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). Discrimination of non-native consonant contrasts varying in perceptual assimilation to the listener's native phonological system. Journal of Acoustical Society of America, 109, 775-794.

Best, C. T., and Strange, W. (1992). Effects of phonological and phonetic factors on cross-language perception on approximants, Journal of Phonetics, 20, 305 – 330.

Booij, G. (1995). The Phonology of Dutch, Oxford University Press, Oxford.

Eimas, P. D., Siqueland, E. R., Jusczyk, P., & Vigorito, J. (1971). Speech perception in infants. Science, 171(968), 303-306.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), Speech perception and language experience: Issues in cross-language research (pp. 233-277). Baltimore: York Press.

Flege, J. (2002). Interactions between the Native and Second-language Phonetic Systems. In P. Burmeister, T. Piske & A. Rohde (Eds) An Integrated View of Language

Development: Papers in Honor of Henning Wode (217-244). Trier: Wissenschaftlicher Verlag.

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer & N. Schiller (Eds) Phonetics and Phonology in Language Comprehension and Production, Differences and Similarities (pp. 319-355). Berlin: Mouton de Gruyter.

Flege, J.E., Takagi, N. & Mann, V. (1995). Japanese adults can learn to produce English /r/ and /l/ accurately. Language & Speech, 38, 25-55.

Goto, H. (1971). Auditory perception by normal Japanese adults of the sounds "L" and "R". Neuropsychologia, 9, 317-323.

Gussenhoven, C. (1995). Illustrations of the IPA: Dutch, in Handbook of the International Phonetic Association, Cambridge University Press, Cambridge, pp. 74-77.

Hamann, S., & Sennema, A. (2005). Voiced labiodental fricatives or glides - all the same to Germans?, in Proceedings of the ISCA Workshop on Plasticity in Speech Perception, London.

Hazan, V. & Barrett, S. (2000). The development of phonemic categorization in children aged 6-12. Journal of Phonetics, 28, 377-396.

Iverson,P., Hazan,V., & Bannister,K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. Journal of the Acoustical Society of America 118, 3267-3278.

Iverson, P., & Kuhl, P. K. (1996). Influences of phonetic identification and category goodness on American listeners' perception of /r/ and /l/. Journal of the Acoustical Society of America 92, 1130-1140.

Iverson,P., Kuhl,P.K., Akahane-Yamada,R., Diesch,E., Tohkura,Y., Kettermann,A., Siebert,C. (2003). A perceptual interference account of acquisition difficulties for non-native phonemes. Cognition 87, B47-B57

Kohler, K. (1999) Illustrations of the IPA: German, Handbook of the International Phonetic Association. Cambridge: Cambridge University Press, 86-89.

Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. Psychometrica, 29, 1-27.

Kuhl, P. K. (2000). A new view of language acquisition. Proceedings of the National Academy of Science, 97, 11850-11857.

Kuhl, P. K. (2004). Early language acquisition: Cracking the speech code. Nature Reviews Neuroscience, 5, 831-843

Kuhl,P.K., Stevens,E., Hayashi,A., Deguchi,T., Kiritani,S., Iverson,P. (2006). Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. Developmental Science, 9, 13-21

Liberman, A., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. Journal of Experimental Psychology, 54, 358-368.

Lively, S.E., Logan, J.S. & Pisoni, D.B. (1993). Training Japanese listeners to identify English /r/ and /l/: II. The role of phonetic environment and talker variability in learning new perceptual categories. Journal of the Acoustical Society of America, 94, 1242-1255.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: a first report. Journal of the Acoustical Society of America, 89, 874-886.

Macmillan, N. A., & Creelman, C. D. (1991). Detection theory : a user's guide. New York: Cambridge University Press.

Moore, B.C.J., Glasberg, B.R. and Baer, T. (1997). A Model for the Prediction of Thresholds, Loudness, and Partial Loudness. Journal of the Audio Engineering Society 45: 224-240.

Miyawaki, K., Strange, W., Verbrugge, R., Liberman, A., Jenkins, J., & Fujimura, O. (1975). An effect of language experience: The discrimination of /r/ and /l/ by native speakers of Japanese and English. Perception & Psychophysics, 18, 331-340.

Näätänen, R. (2001). The perception of speech sounds by the human brain as reflected by the mismatch negativity (MMN) and its magnetic equivalent (MMNm). Psychophysiology, 38, 1-21.

Näätänen, R., Lehtokoski, A., Lennes, M., Cheour, M., Huotilainen, M., Iivonen, A., Vainio, M., Alku, P., Ilmoniemi, R. J., Luuk, A., Allik, J., Sinkkonen, J., & Alho, K. (1997). Language-specific phoneme representations revealed by electric and magnetic brain responses. Nature, 385, 432-434.

Perera, H. S., & Jones, D. (1919). A Colloquial Sinhalese Reader. Manchester, UK: Manchester University Press.

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001) Factors affecting degree of foreign accent in an L2: a review. Journal of Phonetics, 29, 191-215.

Scherer, G. & Wollmann, A. (1985) Englische Phonetik und Phonologie. Berlin: Erich Schmidt Verlag.

Sharma, A., & Dorman, M. F. (2000). Neurophysiologic correlates of cross-language phonetic perception. Journal of the Acoustical Society of America, 107, 2697-2703.

Streeter, L. A. (1976). Language perception of 2-month old infants shows effects of both innate mechanisms and experience. Nature, 259, 39-41.

Trubetzkoy, N. S. (1969). Principles of Phonology (C. A. M. Baltaxe, Trans.). Berkeley: University of California Press. (Original work published 1958).

Werker, J. F., & Yeung, H. H. (2005). Infant speech perception bootstraps word learning, Trends in Cognitive Sciences, 9 , 519-527.

Werker, J.F., & Tees, R.C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. Infant Behavior and Development, 7, 49-63.

Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csepe, V., Ilmoniemi, R. J., & Näätänen, R. (1999). Brain responses reveal the learning of foreign language phonemes. Psychophysiology, 36, 638-642.

Footnotes

1. Although our general view is that it is a labiodental approximant, some descriptions have suggested that it may be bilabial or labiovelar (e.g., Perera and Jones, 1919). We do not know of any systematic phonetic analysis of this phoneme, and it is possible that its articulation may vary in place for different speakers and contexts.

Figure Captions

Figure 1. Synthesized stimulus grid and example spectrograms. The spectrograms display the stimuli at the corners of the stimulus grid, with a duration of 1 second (horizontal axis) and frequencies of 0-5000 Hz (vertical axis). The place dimension varied from labiovelar (low F2) to labiodental (high F2). The manner dimension varied from approximant (high F1, low amplitude of frication, long transition duration after the closure) to voiced fricative (low F1, high amplitude of frication, short transition duration after the closure).

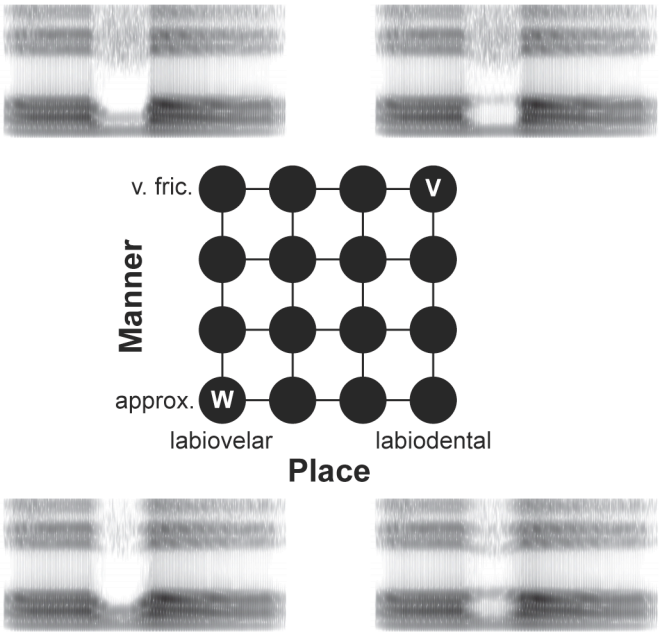
Figure 2. Boxplots (i.e., quartile ranges of individual subject scores) of identification accuracy for natural speech, the degree of contrast between /w/ and /v/ in the subjects' speech, the degree of accent in their production of /w/, and the degree of accent in their production of /v/.

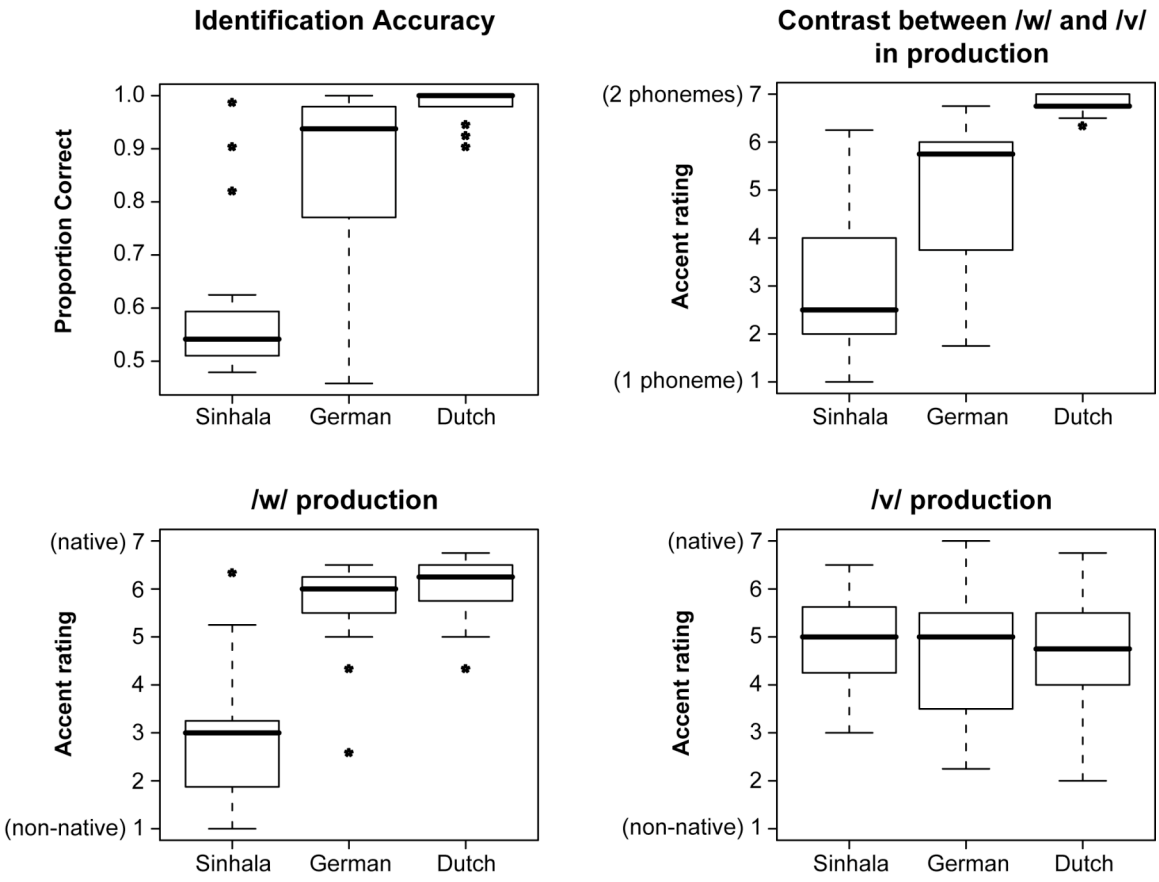
Figure 3. Identification and goodness ratings for synthetic stimuli in terms of the English /w/ and /v/ categories, with the data split according to each subject's language background and whether their identification accuracy for natural English stimuli was greater or less than 75% (Dutch speakers all had higher than 75% accuracy). The grey scale for each stimulus indicates the proportion of identification responses from 100% /w/ (white) to 100% /v/ (black). The size of each circle indicates the average goodness rating for each stimulus, with larger circles for higher goodness ratings. The "W" and "V" labels indicate the corner stimuli that were intended to be the best exemplars of those categories for L1 speakers.

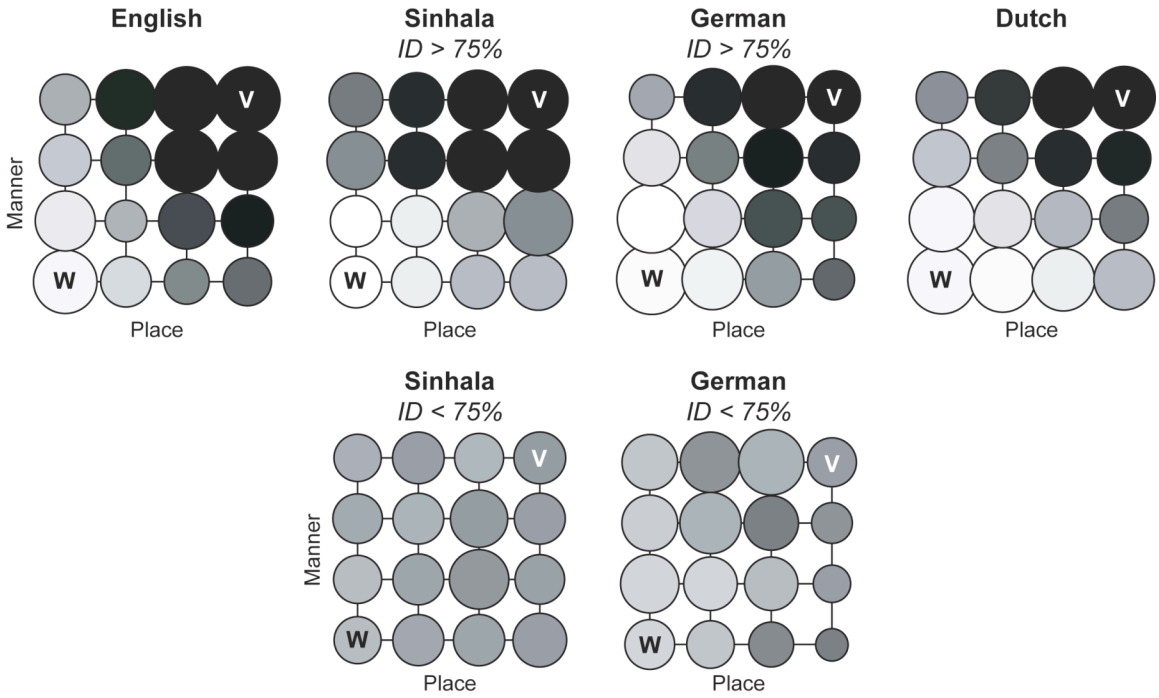
Figure 4. Identification and goodness ratings for synthetic stimuli in terms of each subject's L1, with the data split according to each subject's language background and

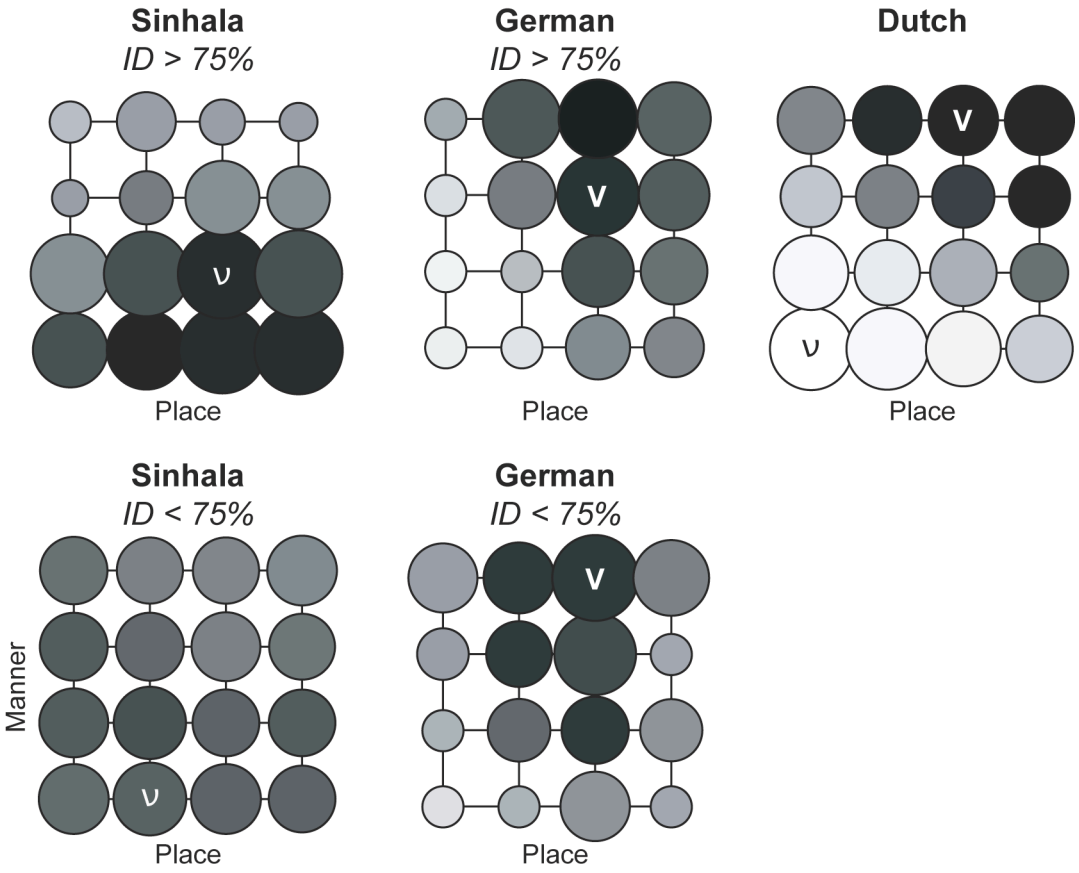
whether their identification accuracy for natural English stimuli was greater or less than 75%. Sinhala speakers identified /ʊ/ or "out of category," German speakers identified /v/ or "out of category," and Dutch speakers identified /ʊ/ or /v/. The grey scale for each stimulus indicates the proportion of identification responses from black for 100% of the closest "v" category (i.e., Sinhala /ʊ/, German /v/, or Dutch /v/) to white for 100% out of category or Dutch /ʊ/. The size of each circle indicates the average goodness rating for each stimulus, with larger circles for higher goodness ratings. The labels indicate the best average stimulus.

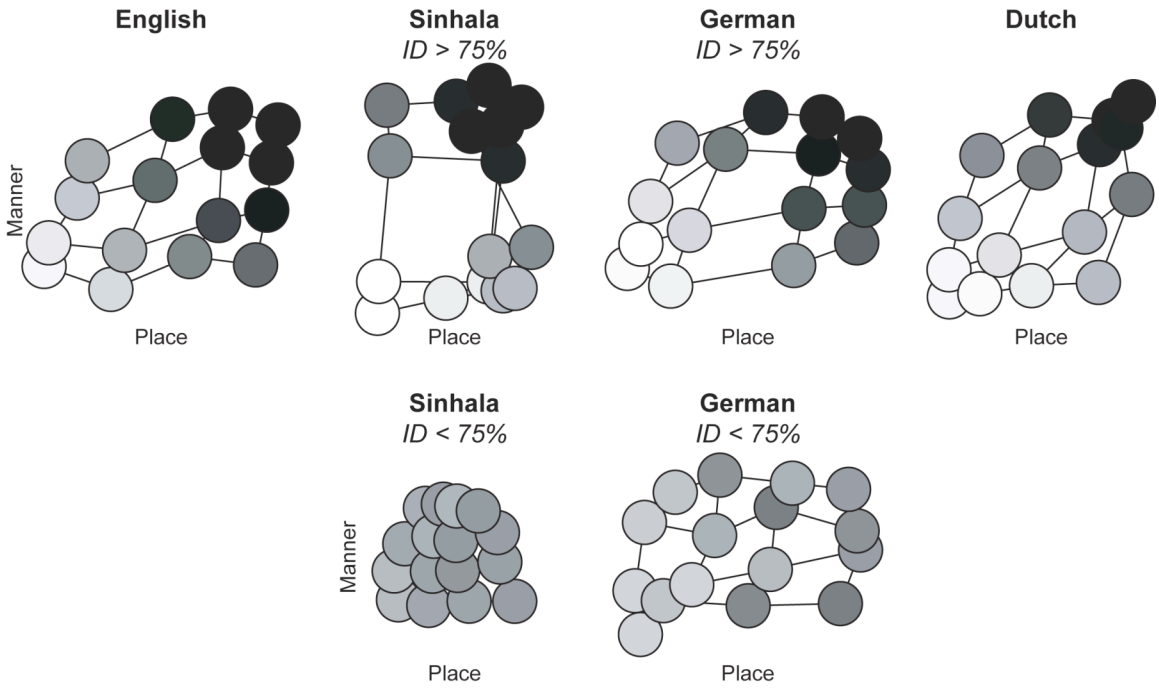
Figure 5. Two-dimensional multidimensional scaling solutions for the synthetic stimuli, with the data split according to each subject's language background and whether their identification accuracy for natural English stimuli was greater or less than 75%. The solutions have been flipped and rotated to match the orientation of the original stimulus grid. The size of each solution was scaled in proportion to their discrimination accuracy. The grey scale for each stimulus indicates the average English identification percentage from 100% /w/ (white) to 100% /v/ (black).


v. fric.
v
Manner
approx.
w
labiovelar
labiodental
Place


Identification Accuracy
Proportion Correct
1.0
0.9
0.8
0.7
0.6
0.5
Sinhala
German
Dutch
Contrast between /w/ and /v/ in production
Accent rating
(2 phonemes) 7
6
5
4
3
2
(1 phoneme) 1
Sinhala
German
Dutch
/w/ production
Accent rating
(native) 7
6
5
4
3
2
(non-native) 1
Sinhala
German
Dutch
/v/ production
Accent rating
(native) 7
6
5
4
3
2
(non-native) 1
Sinhala
German
Dutch


English
Sinhala
ID > 75%
German
ID > 75%
Dutch
Sinhala
ID < 75%
German
ID < 75%
Manner
Place
V
W


Sinhala
ID > 75%
ν
Place
German
ID > 75%
v
Place
Dutch
v
ν
Place
Sinhala
ID < 75%
Manner
ν
Place
German
ID < 75%
v
Place


English
Sinhala
ID > 75%
German
ID > 75%
Dutch
Manner
Place
Place
Place
Place
Sinhala
ID < 75%
German
ID < 75%
Manner
Place
Place