PAM Revisits the Articulatory Organ Hypothesis: Italians' Perception of English Anterior and Nuu-Chah-Nulth Posterior Voiceless Fricatives

Catherine T. Best[1], Cinzia Avesani[2], Michael D. Tyler[1,3] and Mario Vayra[4]

[1] MARCS Institute, Western Sydney University
[2] Consiglio Nazionale delle Ricerche, Italia
[3] School of Social Sciences & Psychology, Western Sydney University
[4] Universitá di Bologna, Italia

Abstract
We perceive non-native speech in terms of similarities to our native phonology, which makes many non-native contrasts difficult to discriminate (e.g., Speech Learning Model [SLM]). However, discrimination is poor mainly when contrasting non-native consonants are both mediocre exemplars of the same native consonant. Discrimination is much better if they are similar to different native consonants, and good if they are nativelike versus deviant exemplars of the same native consonant (Perceptual Assimilation Model [PAM]). The Articulatory Organ Hypothesis (AOH) offers orthogonal predictions that consonants produced by different articulators should be discriminated better than consonants using the same articulator. To compare these models, we tested Italian listeners on non-native English and Nuu-Chah-Nulth fricative contrasts differing in perceptual assimilation, articulatory organs, and articulator use in Italian. Results support PAM and pose challenges for AOH and SLM.

Anne Mette Nyvad, Michaela Hejná, Anders Højen, Anna Bothe Jespersen & Mette Hjortshøj Sørensen (Eds.), ***A Sound Approach to Language Matters – In Honor of Ocke-Schwen Bohn*** (pp. 13-40). Dept. of English, School of Communication & Culture, Aarhus University.

1. Introduction

As adults we apprehend the consonants and vowels in speech with a "native ear." This selective perceptual tuning, shaped by a lifetime of native language (L1) conversational experience, makes comprehending L1 verbal messages largely automatic and fluid, given rapid yet accurate recognition of native spoken words. But this exquisitely supportive tuning of speech perception has a dark side: it leaves us mistuned for reception of the unfamiliar phonetic properties and phonological functions of non-native consonants and vowels, i.e., speech segments that play no role in our own phonological system despite being gainfully employed by other languages. Unsurprisingly, this non-optimal perception of foreign phones hinders second language (L2) speech learning, both for L2 perception and production. And it persists in making verbal comprehension in a later-acquired language slower, more effortful and more easily disrupted than native speech comprehension, even if the listener has become reasonably fluent in the language (see Lecumberri, Cooke, & Cutler, 2010).

Theoretical and empirical investigations into native attunement of speech perception have primarily addressed how experience with a given language or lack thereof influences categorization and discrimination of minimal segmental contrasts, i.e., pairs of consonants or vowels that differ by a single critical phonetic feature that is contrastive in a given language. Moreover, that work has focused largely on "first encounters" of non-native contrasts by listeners naïve to the stimulus language and the target contrasts. However, it is complemented by studies of L2 speech perception by late learners, who come to the task with substantial L1 biases.

In this chapter we compare and contrast three current theoretical models with respect to their hypotheses about the nature of similarities and differences between non-native speech contrasts and those of the listener's native language that shape the perception of non-native speech. We go on to provide findings from a novel study designed to compare those hypotheses, and we discuss the theoretical implications of our findings.

We turn first to the aspects of cross-language perceptual research that are most relevant to those theoretical comparisons and the study we report here. Research and theory on non-native speech perception by naïve adult listeners has from early days focused on their difficulties with categorizing and discriminating minimal phonetic contrasts from unfamiliar languages. The classic proposition that adults possess a native-language *phonological filter* (or sieve) that results in a kind of *phonological deafness* to non-native speech contrasts, as originally posited in the 1930's (Polivanov, 1931/1974; Trubetzkoy, 1939/1969), has been generally accepted based on evidence of naïve listeners' perceptual difficulties with many non-native phonetic contrasts (e.g., Abramson & Lisker, 1970, 1973; Dupoux & Peperkamp, 2002; Iverson et al., 2003; MacKain, Best, & Strange, 1981; Miyawaki et al., 1975; Polka, 1991, 1992; Strange, 1995; Tsao, Liu, & Kuhl, 2006; Werker, Gilbert, Humphrey, & Tees, 1981), as well as of similar difficulties even in early L2 bilinguals (e.g., Sebastian-Gallés & Soto-Faraco, 1999).

However, the stimulus contrasts used in those studies appear to have all been of one particular type, namely cases in which the phonetic characteristics of the contrasting non-native phones align both of them to a single native phoneme. To naïve listeners, these non-native phones are perceived as equally good or poor exemplars of that one native phoneme. Such a narrow range of target stimuli may have led to only partial understanding of the role of experience in non-native speech perception. Findings published since then support that possibility, indicating that non-native phonemes are not all equally difficult to categorize and non-native contrasts are not all equally difficult to discriminate. Performance on both types of tasks is seen to vary when a wider range of non-native phonemes and contrasts has been used.

In light of that variation, several theoretical models of adults' cross-language speech perception have been proposed, which offer a richer, more nuanced view of language-specific perceptual attunement than is captured by the classic phonological deafness concept. We consider here the two models that are most relevant to the perceptual study reported in this chapter[1]: The Perceptual Assimilation Model (PAM: Best, 1995; e.g.,

[1] Other models of cross-language speech perception, while also influential, do not apply as straightforwardly to our reported study on adults' perception of two types of non-native fricative contrasts. Three such models focus on developmental changes in infant rather than adult speech perception as a result of language experience: WRAPSA (Word Recognition And Phonetic Structure Acquisition: Jusczyk, 1993, 1997), NLM (Native Language Magnet: e.g., Kuhl 1993a, b; NLM^{e} = expanded: Kuhl et al., 2008) and PRIMIR (Processing Rich Information from Multidimensional Interactive Representations: Werker & Curtin, 2005; Curtin, Byers-Heinlein & Werker, 2011). Three other models address adult cross-language speech perception more centrally but have focused specifically on vowel perception: NRV (Natural Reference Vowels: Polka & Bohn, 2003, 2011), L2LP (Second Language Linguistic Perception: e.g., Escudero & Boersma, 2004 e.g., Leussen & Escudero, 2015) and ASP (Automatic Selective Perception: Strange, 2011, e.g., Strange & Shafer, 2008).

Best, McRoberts & Sithole, 1988; Best, McRoberts & Goodell, 2001) and the Speech Learning Model (SLM: Flege, 1995, 2003, 2007; e.g., Bohn & Flege, 1990, 1993; Guion, Flege, Akahane-Yamada, & Pruitt, 2000).

PAM was originally created to account for variations in perception across a wider range of types of speech contrasts by listeners of a range of L1s who are completely naïve to the target language and specific contrasts being tested. It has since been extended to address experience-related changes in perception and production of L1, L2 and/or unfamiliar speech contrasts by L2 learners (PAM-L2: Best & Tyler, 2007; e.g., Bundgaard-Nielsen et al., 2011a, b, 2012) and bilinguals (Antoniou et al., 2010, 2011, 2012, 2013; Krebs-Lazendic & Best, 2013). PAM's core principle is *perceptual assimilation*, i.e., the idea that listeners have a strong tendency to perceive unfamiliar non-native phones as exemplars of their L1 phonemes, a tendency grounded in detecting articulatory phonetic and/or phonological similarities to them. If the listener perceives a non-native phone as an acceptable exemplar of a single native phoneme, it is *Categorized*. If a non-native phone is instead perceived to have weaker similarities spread across two or more L1 phonemes, it is an *Uncategorized* consonant or vowel. Very rarely, a non-native phone will fail to be perceived as having any similarity to any native phonemes and will remain *Non-Assimilated*, i.e., be heard as a non-speech sound, as is the case for click-language-naïve English speakers' perception of southern African click consonants (Best, McRoberts & Sithole, 1988; Best, Traill, Harrison, Carter, & Faber, 2003).

When two contrasting non-native phones are each categorized to a different L1 phoneme, this constitutes *Two Category* (TC) assimilation, and discrimination is predicted to be excellent. If instead the members of a non-native contrast are both categorized to the same single L1 phoneme, they may be perceived as equally good or poor exemplars of it (*Single Category* assimilation: SC) or one may be a perceptibly poorer fit than the other (a *Category Goodness* difference in assimilation: CG). Discrimination of CG contrasts is predicted to be very good but significantly lower than for TC contrasts, whereas SC contrasts are predicted to be poorly discriminated. If one or both members of a non-native contrast are uncategorized (UC or UU, respectively), discrimination performance level will depend on the subtype of uncategorized assimilation(s) involved, for example, whether or not the contrasting non-native phones show overlap in the L1 phonemes to which similarities are perceived (see Faris, Best, & Tyler, 2016; Faris, Tyler, & Best, 2018). PAM-L2 (Best & Tyler, 2007) makes the case that L2 learning is most likely to result in improved categorization and discrimination of

L1 contrasts that were initially CG or uncategorized assimilations. Note that discrimination of non-native contrasts is now assumed to be better for non-overlapping than overlapping assimilations, within each of the relevant contrast assimilation types: TC, UC and UU (see Faris, Best, & Tyler, 2016; Fenwick, Best, David, & Tyler, 2017; Tyler, Best, Faber, & Levitt, 2014).

Whereas PAM's central aim is to account for variations in non-native speech perception, SLM instead aims to understand the factors that give rise to foreign accent in L2 speech production. Still, SLM makes strong perceptual assumptions, arguing that the most important source of foreign accent is L1 biases in the speaker's perception of L2 speech. A core SLM premise is that L2 speech production can only be as accurate as L2 speech perception permits. SLM posits that L1 perceptual biases lead to *equivalence classification* of L2 phones as being either *identical*, or *similar*, or *new* with respect to L1 phonemes. *Identical* L2 phones pose no difficulty for perception or production, of course, as they correspond well to L1 phonemes. And although *new* L2 phones may pose some difficulties initially, the model predicts that they will be fairly easily established as new, separate L2 categories in both perception and production. In contrast, SLM predicts that equivalence classification of *similar* L2 phones to L1 phonemes results in a persisting L1 perception bias and L1-accented production.

Thus, PAM and SLM have somewhat different yet overlapping and/or complementary foci and conceptual principles. While PAM's central goal is to account for how L1 experience shapes speech perception, particularly of non-native minimal consonant contrasts by naïve listeners, SLM's is to understand the factors contributing to accented speech production by L2 learners/bilinguals, with particular focus on vowels as individual categories. Nonetheless, although the primary foci of the two models are complementary, their secondary emphases bring them back to common ground. SLM considers L1 influences on non-native speech *perception* to be the most important contributor to accented L2 speech production and has fostered investigations of L2 perception (e.g., Bohn & Flege, 1990). Conversely, PAM has been extended to address speech *production* (e.g., Antoniou, Best, Tyler, & Kroos, 2010, 2011; Bundgaard-Nielsen, Best, Kroos, & Tyler, 2012) as well as perception (e.g., Antoniou et al., 2012, 2013; Bundgaard-Nielsen et al., 2011a, b; Krebs-Lazendic, & Best, 2013). In theoretical terms, both PAM and SLM posit that non-native speech is perceived in relation to native (L1) phonemes. Moreover, it is far from obvious whether their proposed processes of perceptual assimilation and equivalence classification, respectively, differ conceptually very much if at all.

The models do differ, however, in their assumptions about the nature of speech information that perceivers use in the L2aL1 process. PAM posits that the process relies on perceiving information about the articulatory gestures that produced the phones, whereas SLM assumes that it relies on acoustic-phonetic similarities between L2 and L1 phones. Other points of relative difference are that SLM investigations have focused more on vowel than consonant perception, and on individual phonetic rather than on minimal contrasts, whereas PAM research has specifically addressed contrasts and has examined consonant perception more than vowels. Nonetheless, some SLM studies have examined consonants (e.g., Bohn & Flege, 1993), while some PAM studies have addressed perception of vowel contrasts, both from other languages (Bundgaard-Nielsen et al., 2011a, b; Faris, Best, & Tyler, 2016, 2018; Tyler, Best, Levitt, & Faber, 2014) and from other L1 regional accents (Best et al., 2013, 2015a, b; Shaw et al., 2014, 2018). PAM has also been applied to perception of non-native lexical tone contrasts (Hallé, Chang, & Best, 2004; Reid et al., 2015; So & Best, 2010, 2011, 2014).

Consideration of lexical tone perception by naïve listeners of non-tone L1s raises an important question that has not been directly addressed by either PAM or SLM: How might perceptual assimilation/classification work in cases where the non-native contrast uses articulatory/acoustic-phonetic properties that are not employed for segmental contrasts in the listeners' L1? This is one of the questions addressed in the study we report in this chapter. Whereas tone languages engage laryngeal mechanisms to produce fundamental frequency (and sometime voice quality) differences that serve as sub-lexical phonological contrasts that are analogous to minimal segmental contrasts between consonants or vowels, non-tone languages only use tonal patterns at higher, suprasegmental prosodic levels in the phonological hierarchy (e.g., stress, accent, and phonological and intonational phrase boundaries). This means that non-tone language speakers cannot assimilate non-native lexical tones to L1 segments; they may instead perceive them in relation to higher-level prosodic patterns in their L1. Neither SLM nor PAM were designed to address this type of phonological tier discrepancy between the non-native target items and the most likely L1 referent categories (see Best, 2019). Indeed, the phonological tier mismatch is reflected in the performance of naïve non-tone L1 listeners on PAM-based perceptual tests with non-native lexical tone contrasts, where their assimilations to L1 prosodic categories have been fairly weak while conversely their discrimination of tone pairs has been better than expected from those assimilation patterns (Hallé, Chang, & Best, 2004; Reid et al., 2015; So & Best, 2010, 2011, 2014).

But what might happen when there is *not* a phonological tier mismatch? How might listeners perceive non-native segmental contrasts, e.g., consonants, that use articulatory/acoustic-phonetic properties not employed in their L1 phonology at *either* the segmental or suprasegmental level? If we extrapolate from SLM principles, it seems likely that such consonants would not be equivalence classified as either *identical* or as *similar* to even the acoustically closest L1 consonants because they would nonetheless be too distant from all native consonants; they would instead be perceived as *new* consonants. While they should therefore be easily distinguished from any L1 consonants, it is not clear from SLM whether contrasting pairs of such non-native consonants would be easily discriminated from *each other*, because its principles focus on individual L2 phones in relation to L1 phonemes, not on discrimination of L2 contrasts. And although PAM directly addresses perception of non-native contrasts, it has not explicitly considered how the assimilation may be affected when the articulators involved in the non-native contrast are not employed in the L1. Will such non-native phones be categorised to the most articulatorily similar L1 consonants, or instead more ambiguously assimilated as uncategorized consonants (or possibly even Non-Assimilated, i.e., heard as nonspeech)? And how would discrimination be expected to be affected by these differing possibilities? These questions are examined by the study we report here.

The Articulatory Organ Hypothesis of infant speech perception (AOH: Goldstein & Fowler, 2003; see also Best, Goldstein, Tyler, & Nam, 2016) could potentially offer some more straightforward predictions, however, if we extend it to adult non-native consonant perception. Originally designed to predict developmental changes in infants' perception of native and non-native phonetic contrasts as a result of experience (Studdert-Kennedy & Goldstein, 2003; see Best & McRoberts, 2003), the AOH posits that between-organ articulatory contrasts are easy to distinguish perceptually, even if they are non-native (do not occur in the infant's environment), whereas within-organ contrasts are more difficult to discriminate and to learn even if they occur in native speech. In between-organ contrasts the contrasting consonants use different primary articulators, e.g., the ejective stops of Tigrinya, /p’/ (lips) versus /t’/ (tongue tip), whereas in within-organ contrasts the consonants use the same primary articulator but with contrasting place, manner or voicing, e.g., the Hindi dental versus retroflex coronal stops /d̪/-/ɖ/ (tongue tip for both, but at two contrasting places). A few speech perception studies have tested the AOH with infants, with mixed results (*supported*: Best & McRoberts, 2003; *compatible*: Kuhl et al., 2006; Polka, Colantonio, & Sundara, 2001; *not supported*: Tyler, Best, Goldstein, & Antoniou, 2014). Adults, however, with their much greater L1 experience, might possibly show more clearly differentiated perceptual responses to non-native within- versus between-organ contrasts, especially for articulatory organs not employed distinctively in their L1.

To examine the sets of questions raised above, a listener language and target stimulus languages were needed for which one set of non-native consonant contrasts uses articulatory organs employed in the listeners’ native language, while another set uses articulatory organs not employed in their language. Italian meets these requirements with respect to two sets of non-native voiceless fricative place of articulation distinctions. Regarding the native-articulators set, Standard Italian has a series of place contrasts among anterior voiceless fricatives that employ the lips (labiodental /f/) and the tongue tip (lamino-alveolar or dental /s/ [respectively, Mioni, 2001, p. 157; Bertinetto & Loporcaro, 2005], and palato-alveolar /ʃ/, which also has secondary tongue body and lip constriction). English offers a set of voiceless anterior fricatives using the lips and/or tongue tip, /f, θ, s, ʃ/, which adds an interdental place of articulation for tongue tip constrictions that is lacking in Italian (no /θ/). Thus, the English series offers two non-native contrasts for which the primary articulators are nonetheless used in Italian, /f/-/θ/ and /θ/-/s/. Those two pairs also provide the required between-organ (/f/-/θ/: lips vs. tongue tip) versus within-organ contrasts (/θ/-/s/: both tongue tip). The remaining minimal-place contrast (/s/-/ʃ/) we define here as a mixed/overlapping organ contrast, given that both of these fricatives use tongue tip but only /ʃ/ also involves constriction of tongue body and lips.

For the non-native-articulators set, that is, consonants that use articulators not employed in Italian, we chose the Nuu-Chah-Nulth (a First Nations Wakashan language, British Columbia, Canada) four-way series of posterior voiceless fricatives, /x, χ, ħ, h/ (velar, uvular, pharyngeal, glottal), in which the primary articulatory organs are either not used at all in Italian phonological contrasts or are not used for fricative manner. These posterior fricatives provide three non-native minimal-place contrasts. Within-organ /x/-/χ/ both use the tongue body, which is only employed in Italian for velar stops, not fricatives; Italian does not employ the uvular place for any consonants. Nuu-Chah-Nulth /ħ/-/h/ is a between-organ contrast, for which /ħ/ employs the tongue body plus tongue root (see Carlson & Esling, 2003) while the articulator for /h/ is the glottis (vocal cords). Italian does not use the tongue root (pharyngeal constriction) for consonant contrasts, nor does it employ the glottis as an active articulator for fricatives, for which it does not have voicing distinctions. The final minimal-place contrast, /χ/-/ħ/, is mixed/overlapping organ (tongue body vs tongue body+root) and involves places of articulation and an organ (tongue root) that are not employed in Italian, as well as the fricative manner that is not used in Italian posterior to the hard palate.

We examined native Italian listeners' assimilation of these English and Nuu-Chah-Nulth fricatives to Italian consonants, using an L1-categorization and goodness rating task. An AXB discrimination task was used to assess their discrimination of the three English and three Nuu-Chah-Nulth minimal-place distinctions. In order to avoid having the L1 categorizations contaminate discrimination performance, for each stimulus language the AXB task was completed first, followed by the categorization and rating task.

2. Method

2.1 Participants. The listeners were 24 native speakers of Italian who also spoke Veneto Dialect (Venetan); all studied/worked at the University of Padova (M_{age}=27.96 years; range=19-43; 13 female). Only one participant had ever lived outside of Veneto (one year each in Florence and Stockholm, during his late 30's). All had either acquired both languages from birth (*n*=17) or had acquired Italian first and Venetan as an early second language (*n*=7). They gave high self-ratings for comprehending (*M*=4.83 on a 5-point scale) and speaking (*M*=4.5) Venetan. Twenty-two learned Central Venetan, which like Italian has no interdental consonants. The other two had learned Northeast Venetan (Treviso), which has an interdental fricative /θ/ (Zamboni, 1974, 1988; see also Avesani, Galatà, Vayra, Best, Di Biase, Tordini, & Tisato, 2016; Avesani, Galatà, Best, Vayra, & Ardolino, 2017). For one of these two participants, Italian was the native language, Venetan was later-learned and weaker; his North Venetan experience did not enhance his detection of the dental feature of English /θ/, which he categorised 80%

of the time as /f/. For the other, Italian and Venetan were learned from birth and were equally strong; nonetheless he categorized English /θ/ similarly to the majority, as a mediocre Italian /t/ (see 3. Results).

Only one participant had not learned any additional languages at school. Twenty-two had studied English ($M_{\text{onset-age}}$=7.5 years, range=4-11; M_{duration}=11.5 years, range=7-20). Although this suggests they should be familiar with English /θ/ and /h/, we note that their mean self-ratings for speaking (*M*=3.0 out of 5) and comprehending English (*M*=3.3) were only fair. Other foreign languages were learned by fewer people and received even lower self-ratings: Spanish (*n*=6, M_{speak}=2.2; $M_{\text{comprehend}}$=2.2), French (*n*=12, M_{speak}=1.9; $M_{\text{comprehend}}$=2.1) and German (*n*=5, M_{speak}=2.6; $M_{\text{comprehend}}$=2.4). These languages do have some posterior fricatives, though none has the full array found in Nuu-Chah-Nulth. Spanish has only /x/, with some uvular variants [χ] in northern and central Spain (Hammond, 2001). Standard French has only [χ] as a positional devoiced allophone of its voiced uvular fricative /r/ ([ʁ]) following voiceless stops (e.g., *lettre*). German has three voiceless guttural fricatives: /x/, /χ/ and /h/ (no pharyngeal /ħ/), with /x/ displaying two vowel-context conditioned allophones, palatal [iç] and velar [ax̠].

2.2 Stimuli

2.2.1 English. Multiple tokens of the English anterior voiceless fricatives labiodental /f/, interdental /θ/, alveolar /s/ and palato-alveolar /ʃ/ (*n*=12 each) in /Ca/ syllables were recorded in random order at Western Sydney University, Australia, by an Australian female speaker in her late 50's whose voice quality was similar to that of the Nuu-Chah-Nulth speaker (see 2.2.2). To ensure that discrimination of the English and Nuu-Chah-Nulth fricatives contrasts would not be confounded by non-criterial acoustic differences between the stimuli of the two languages, the English tokens were adjusted in Praat (Boersma & Weenink, 2009), using overlap-add resynthesis, to achieve a similar mean and range of consonant and vowel durations as the Nuu-Chah-Nulth stimuli. To reduce the possibility that /θ/-/s/ would be discriminated solely on intensity differences, /f/ and /θ/ were additionally amplified by 5 dB and /s/ reduced by 6 dB. This still left /s/ with a higher amplitude than /f/ and /θ/ to maintain naturalness. Vowel intensities for all tokens were adjusted to the same level of acoustic intensity.

2.2.2 Nuu-Chah-Nulth. Multiple natural tokens of the Nuu-Chah-Nulth posterior voiceless fricatives velar /x/, uvular /χ/, pharyngeal /ħ/ and glottal /h/ (*n*=15+ each) in /Caː/ syllable context (e.g., /xaː/) were produced in random order by a female native speaker in her 60's from the traditional tribal area on Vancouver Island, British Columbia, Canada. The recordings were made at the University of British Columbia. We chose a speaker of the elder generation because they maintain the /x/-/χ/ distinction, which younger speakers may have lost, i.e., /x/-/χ/ appears to have undergone merger over recent decades. For the perceptual tests we selected four tokens of each of the four target syllables, matched across fricative categories in duration, amplitude and pitch contour. Vowel intensities of all tokens were adjusted to the same level of acoustic intensity. (Note: the /x, χ, ħ/ and /f, θ, s/ stimuli were used in studies with infants in Tyler, Best, Goldstein, & Antoniou, 2014.)

2.3 Procedure. Participants completed a discrimination test followed by a categorization and rating test on the Nuu-Chah-Nulth consonants with respect to an on-screen array of Italian consonant choices, then discrimination followed by categorization and ratings of the English consonants. Discrimination was assessed prior to categorization in order to minimize confounding effects of prior categorizations on discrimination.

2.3.1 Discrimination. A categorial AXB discrimination task was used because it has lower memory demands and minimizes response biases relative to other standard discrimination protocols (see Best & Strange, 1992; Pollack & Pisoni, 1971; Strange & Shafer, 2008). On each trial participants received three stimuli separated by 1 s interstimulus intervals (ISIs), of which the first and third (A and B) were contrasting consonants and the middle item, X, was a different token of either the A or B consonant category. They had to indicate as quickly and accurately as possible whether X matched category A or B. Each stimulus triad appeared in four trial configurations: AAB, ABB, BBA, BAA. Inter-trial intervals (ITIs) were 3.5 s. Three contrasts were tested for each language, with 48 trials per contrast (4 trial types x 4 stimulus token triads x 3 repetitions) in separate blocks: Nuu-Chah-Nulth /x/-/χ/, /χ/-/ħ/ and /ħ/-/h/ and English /f/-/θ/, /θ/-/s/, /s/-/ʃ/. Test order of the contrast blocks within each language were randomized across participants. Before the first discrimination block they received a short set of practice AXB trials that used an unrelated non-native lateral fricative voicing distinction from isiZulu (from Best, McRoberts, & Goodell, 2001).

2.3.2 Categorization and goodness ratings. On each trial of the categorization task following discrimination in each language, participants were presented with a single token and had to indicate which Italian consonant the non-native token sounded most similar to, selecting from a set of printed on-screen consonant+/a/ syllables using standard Italian spelling, which transparently conveys to Italians how to pronounce the consonants: FA, SA, SCIA, PA, TA, CA, LA, GLIA, RA, UA, CIA, JA, ZA and HA. We also provided examples of Italian words beginning with the relevant consonant. The fine-grained pronunciations of the initial consonants in the Venetan variety of Italian spoken by our participants are given in narrow IPA transcription as follows: FA [f], SA [s], SCIA [ʃ], are pronounced as in English. The voiceless stops differ from English, however, as they are unaspirated, which in initial position is phonetically more similar to English *voiced* stops: PA [p], TA [t], CA [k]. The glides of Italian also differ in phonetic details from English: LA [l] is lighter than in English (flatter tongue, less velarised), RA [r] is an alveolar tap/trill, GLIA [ʎa] is a palatal lateral, and UA [ua] differs dynamically from English [w]. The Italian affricate CIA is pronounced like English <ch> [tʃ]. The spelling JA was taken from the English loanword <jazz> because Italians pronounce it as [dʒa], whereas the Italian spelling GIA would be pronounced as a bisyllable [dʒia]. ZA [dza] is a dental affricate that does not exist in English. HA is pronounced [_a] in Italian, i.e., with a "silent h" [∅][2] rather than an aspirated [h] preceding the vowel, which would likely have a glottal stop onset. We asked them to choose the item with the most similar pronunciation of the consonant in Italian. Given that 22 of our 24 participants had studied English, we cannot rule out the possibility that some may have used <H> to indicate the English aspirated glottal fricative [h] despite our instructions to focus on Italian pronunciations. However, as a reminder, they self-rated their proficiency in speaking and understanding spoken English to be mediocre on average.

2 <H> in Italian spelling is an orthographic convention. If it is inserted between <C, G> and <I, E>, it specifies that <C, G> are pronounced as the stops [k] and [g], rather than as the palatalized affricates [tʃ] and [dʒ] that are indicated by <CI, CE> and <GI, GE>, i.e., with no <H> intervening. Initial <H> also occurs but is silent in the written first, second and third person singular and third person plural forms of the verb <AVERE> 'to have' (HO, HAI, HA and HANNO, respectively).

After making their choice, they heard the same token again and had to rate how good a match it was to their selected Italian consonant, using a 1-7 Likert scale (1 = poor match, 7 = excellent match). There were 64 categorization trials per language (4 target consonants x 4 tokens each x 4 repetitions of the set), presented in random order. The first categorization test was preceded by a short practice set of the Zulu voiced and voiceless lateral fricatives.

3. Results

3.1 Categorization and goodness ratings. Although the categorization test was run *after* the discrimination test for each language, the categorization results will be presented first, as they determine the assimilations of the non-native fricatives to Italian consonants, which in turn provides the PAM predictions for discrimination performance differences among contrasts.

Italian labels	**NON-NATIVE TARGET FRICATIVES**							
	English				**Nuu-Chah-Nulth**			
[IPA]	**/f/**	**/θ/**	**/s/**	**/ʃ/**	**/x/**	**/χ/**	**/ħ/**	**/h/**
<C> [k]					*24* (3.37)			
<F> [f]	**96** (6.32)	*25* (4.62)						
<H> [∅]					**57** (2.59)	**84** (3.98)	**95** (5.38)	**93** (5.20)
<J> [dʒ]					*15* (3.69)			
<S> [s]			**96** (6.20)					
<SCIA> [ʃ]				**96** (6.27)				
<T> [t]		**66** (3.08)						

Table 1. Mean percent categorizations and goodness ratings (1-7 scale; in parentheses) of each English and Nuu-Chah-Nulth target fricative to the Italian consonant choices (in Italian orthography and IPA). Boldface indicates significantly above chance. Italicized indicates significantly above chance but chosen significantly less often than the modal choice. Only labels selected significantly above chance (7%) are displayed. Italian labels chosen < 7% of the time for any target: <CIA>, <LA>, <PA>, <RA>, <ZA>, <GLIA>, <UA>).

Table 1 shows the categorization and goodness ratings for each English and Nuu-Chah-Nulth target fricative in relation to the Italian consonant choices. We used statistical criteria, rather than a pre-set threshold as in previous research, to determine whether a target consonant was Categorized to a single native consonant or was instead Uncategorized. The thresholds used in prior studies have not been standardized (varying among 50%, 70%, 90%), and their rationales have been somewhat subjective and arbitrary, which has made cross-study comparisons problematic. To address this, we created a new statistical criterion that can be applied systematically across different types of targets (consonants, vowels, tones) and across studies. Specifically, we designate a non-native target as Categorized if one L1 consonant was chosen significantly more than all other choices, and if it was also chosen significantly above chance.[3] If it did not meet both criteria it was deemed Uncategorized.

By these criteria, all English fricatives were Categorized. Although <T> was chosen for English /θ/ only 66% of the time on average, this was significantly above chance and significantly greater than choices of the next highest Italian category, <F> (*M*=25%), $t_{(23)}$=2.76, *p*<.015. Each English fricative was Categorised to a different Italian label, making all pairwise assimilations Two Category (TC) contrasts. However, the TC assimilation for /f/-/θ/ differed in fine-grained detail from that for /f/-/s/ and /s/-/ʃ/. There was partial overlap in the use of <F> for both English /f/ and /θ/, yielding an overlapping TC assimilation pattern, or TC-O (see Tyler, Best, Faber, & Levitt, 2014), whereas there was no overlap in choices for the other two English contrasts, which were therefore non-overlapping TC assimilations (TC-N). In light of our previous arguments that overlapping assimilations should be more difficult to discriminate than non-overlapping ones within a contrast assimilation type (Faris, Best, & Tyler, 2016; Fenwick, Best, David, & Tyler, 2017; Tyler, Best, Faber, & Levitt, 2014), /f/-/θ/ should show poorer discrimination than /θ/-/s/ and /s/-/ʃ/, which should not differ from each other.

3 These criteria differ from those proposed in Faris, Best, & Tyler (2016), which apply to a non-native item (vowels in their study) that had first been designated as Uncategorized according to a 50% threshold, i.e., the top choice native category was chosen less than half the time. If a single <50% category was nonetheless significantly above chance and no other categories were significantly above chance it was considered Focalised-Uncategorised. By our new statistical criteria, none of the English or Nuu-Chah-Nulth fricatives were Uncategorized.

In contradistinction to the English fricatives, a single label, <H>, was the most common choice for all four Nuu-Chah-Nulth fricatives; again they all met the Categorized criteria. For /x/, <H> was chosen ($M = 59\%$) significantly more often than the two next-higher, above-chance category choices of <C> (M=23%), $t_{(23)}=2.764$, $p<.012$, and <J> (M=15%), $t_{(23)}=3.69$, $p<.002$. The pairwise assimilation patterns for the Nuu-Chah-Nulth contrasts were constrained both by the categorization of all four fricatives to <H> and by the more dispersed choices for /x/. Contrast /χ/-/ħ/ was assimilated to Italian <H> as a Category Goodness difference (CG) contrast, given that the ratings of goodness of fit to <H> were significantly lower for /χ/ (M_{rating}=4.0) than /ħ/ (M_{rating}=5.4), $t_{(23)}=4.58$, $p<.001$. The /x/-/χ/ contrast was also a CG assimilation, in which /x/ was rated a significantly poorer <H> (M_{rating}=2.6) than /χ/ was (M_{rating}=4.0), $t_{(23)}=3.35$, $p<.002$. However, /ħ/ and /h/ ratings as <H> did not differ significantly, making the assimilation of /ħ/-/h/ a Single Category (SC) contrast. Based on PAM predictions (Best, 1995), then, /ħ/-/h/ should show poorer discrimination than /χ/-/ħ/ and /x/-/χ/, which should not differ from each other, yet should show lower discrimination than the English TC contrasts.

3.2 Discrimination. Discrimination was above chance (50% on AXB tasks) for each of the six contrasts. Five contrasts were significantly above chance at $p<.001$: /f/-/θ/, $t_{(23)}=9.418$; /θ/-/s/, $t_{(23)}=11.988$; /s/-/ʃ/, $t_{(23)}=16.522$; /x/-/χ/, $t_{(23)}=8.770$; /χ/-/ħ/, $t_{(23)}=6.367$. Nuu-Chah-Nulth /ħ/-/h/ performance was also above chance, $t_{(23)}=2.326$, $p<.03$.

We conducted a repeated measures Analysis of Variance (ANOVA) on the accuracy data for four within-subjects factors: Language (English, Nuu-Chah-Nulth), Contrast Type (between-organ, within-organ, mixed/overlapping organs), Consonant (whether X in the AXB trials was the more anterior or more posterior consonant of the contrast) and Match (whether X matched the consonant of the first or third item in the AXB trial). For English, the between-organ contrast was /f/-/θ/ (lips vs. tongue tip constriction), the within-organ contrast was /θ/-/s/ (both tongue tip constrictions) and the mixed/overlapping organ contrast was /s/-/ʃ/ (tongue tip vs. tongue tip+body constrictions). For Nuu-Chah-Nulth the between-organ contrast was /ħ/-/h/ (tongue body+root constriction vs. glottis wide for aspiration), the within-organ contrast /x/-/χ/ (both tongue body constrictions) and the mixed/overlapping organ contrast /χ/-/ħ/ (tongue body vs. tongue body+root constriction). The results are displayed in Figure 1.

The main effect of Language was significant, $F_{(1,23)}$=75.31, p<.001, η_p^2=.766, indicating that mean discrimination accuracy was significantly higher overall for English (M=84.9%) than for Nuu-Chah-Nulth (M=63.4%). Contrast Type was also significant, $F_{(2,46)}$=31.84, p<.001, η_p^2=.581. Pairwise tests on this effect indicated that, counter to AOH predictions, performance was significantly *lower* rather than higher for the between-organ contrasts (M=66.1%) relative to both the within-organ (M=78.6%), p<.001, and mixed/overlapping contrasts (M=77.7%), p<.001, which did not differ significantly. However, these main effects were modulated by three significant interactions: Language x Contrast Type, $F_{(2,46)}$=4.42, p=.018, η_p^2=.161; Contrast Type x Consonant, $F_{(2,46)}$=5.34, p=.008, η_p^2=.188; and Language x Contrast Type x Consonant, $F_{(2,46)}$=4.38, p=.018, η_p^2=.161.

To break down the three-way interaction, we ran separate ANOVAs for each language on the within-subjects factors Contrast Type x Consonant x Match. For the English ANOVA, only the main effect of Contrast Type was significant, $F_{(2,46)}$=12.051, p<.001, η_p^2=.344. Pairwise comparisons indicated that performance on the between-organ contrast, /f/-/θ/, was significantly lower (M=78.6%) than on the within-organ /θ/-/s/ (86.5%), p<.007, and the mixed/overlapping /s/-/ʃ/ contrast (M=89.6%), p<.001, but the latter two contrasts did not differ from each other.

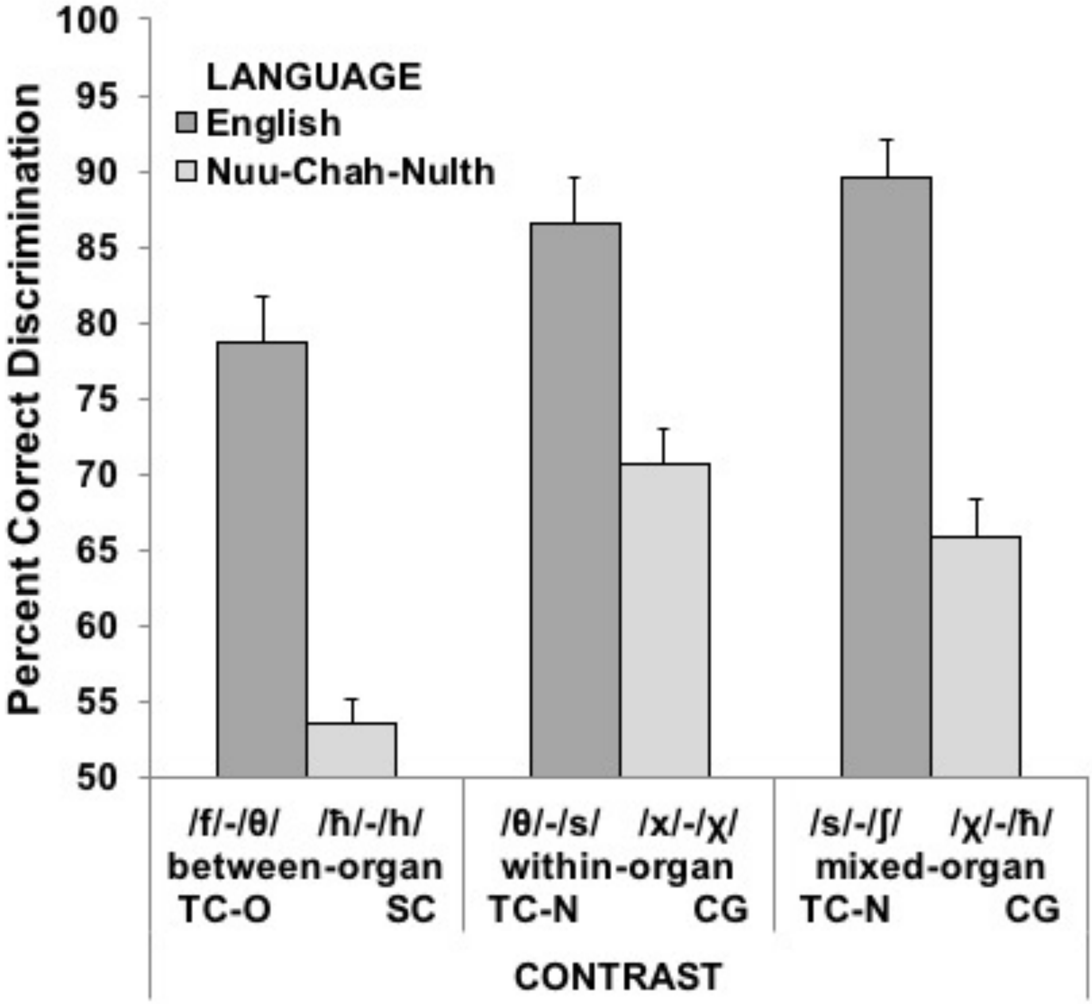

Figure 1. Mean percent correct discrimination of the English and Nuu-Chah-Nulth voiceless fricative contrasts, with Contrast Type and assimilation pattern (from the Categorization/rating results) displayed beneath each discrimination pair on the x-axis.

The significantly lower performance on /f/-/θ/, and the split in categorizations of /θ/ as <T> versus as <F> (see Table 1), are consistent with our predictions regarding TC-O (overlapping) assimilation for this contrast. However, we noted individual variability in the tendency to report <F> for /θ/, which led us to examine individual participants' categorizations of /θ/. In total, 17 of the 24 participants (71%) selected <T> more than 50% of the time. Of the other seven participants (29%), six selected <F> more than 50% of the time. The remaining participant did not select any one label more than 50% of the time, but her highest response was <F> (31%) and she never selected <T>, so we grouped her with the six who had categorized /θ/ to <F>. Note that the /f/-/θ/ contrast is a non-overlapping Two Category (TC-N) assimilation for listeners who categorize /θ/ to Italian <T>, but a Category Goodness difference (CG) assimilation for those who categorized /θ/ as Italian <F>; they gave a very good rating of English /f/ as <F> (M_{rating}=6.04) as compared to a moderate rating of /θ/ as <F> (M_{rating}=4.91), $t_{(6)}$= .29, p<0.032 (one-tailed, as better ratings as <F> are predicted for /f/ than /θ/ stimuli). Given these subgroup differences, we conducted a new ANOVA on the English discrimination data, with Subgroup (/θ/-as-<T> vs. /θ/-as-<F>) as a between-subjects factor, and Contrast Type (between-organ, within-organ, mixed/overlapping organs) as a within-subjects factor. Only the interaction was significant, $F(2, 44)$ =5.60, p=.007, η_p^2=.20. Post-hoc *t*-tests revealed a significant difference between the Subgroups on discrimination of /f/-/θ/, $t(22)$=2.81, p=.01, with the TC-N /θ/-as-<T> categorizers showing better discrimination (M=83%) than the CG /θ/-as-<F> categorizers (M=67%). There was no Subgroup difference for the other two contrasts (see Figure 2).

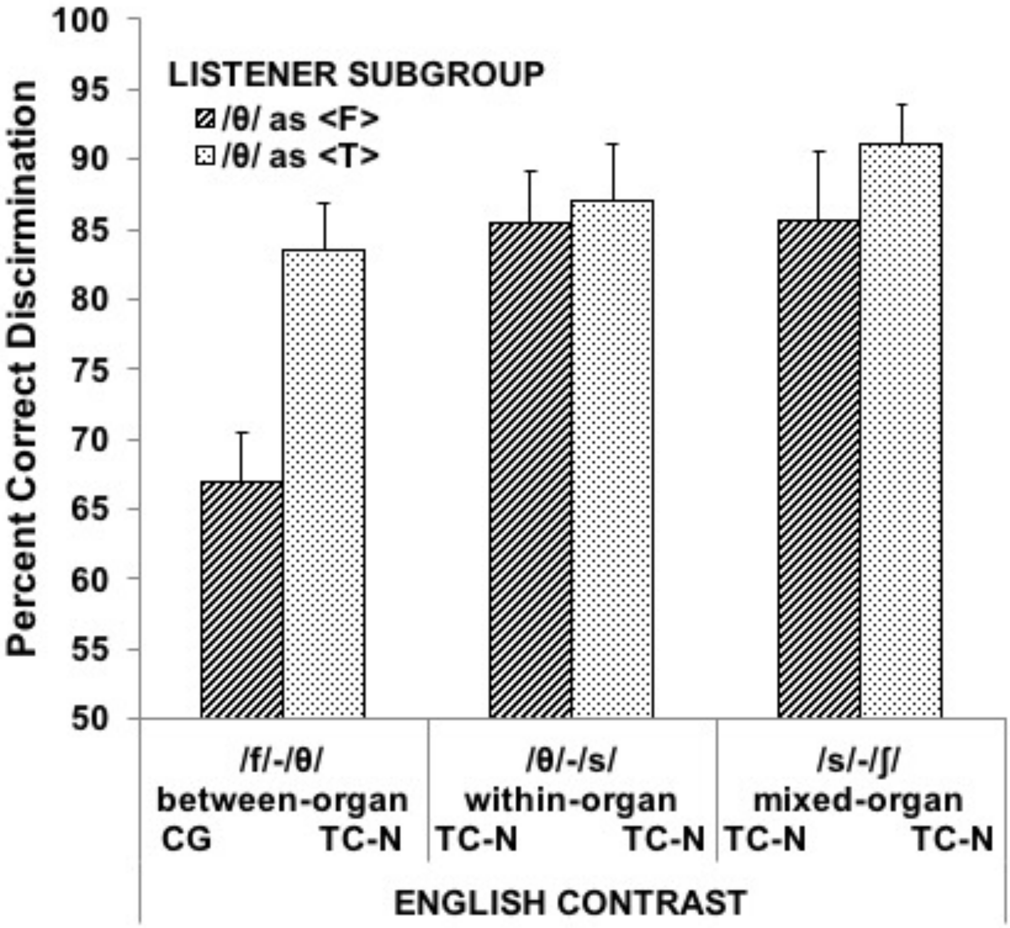

Figure 2. Mean percent correct discrimination of the English voiceless fricative contrasts by the participants who assimilated English /θ/ as <T> versus the participants who assimilated it as <F>.

The Nuu-Chah-Nulth breakdown ANOVA also found a significant main effect of Contrast Type, $F_{(2,46)}$=24.5, p<.0001, η^2=.516, for which pairwise tests indicate that discrimination of the between-organ SC contrast, /ħ/-/h/, was significantly lower (M = 53.6%) than the within-organ /x/-/χ/ (M=70.7%), p<.001, and mixed/overlapping /χ/-/ħ/ CG contrasts (M=65.9%), p<.001, which did not differ significantly. The significant Contrast Type x Consonant interaction, $F_{(2,46)}$=6.68, p<.003, η^2=.225, revealed that discrimination of the CG within-organ contrast /x/-/χ/ was better when X in the AXB trials was /x/ (M=76%) than when it was /χ/ (M=65.3%), but there was no Consonant effect for the SC between-organ contrast /ħ/-/h/ (M=50.1 vs. 56.4%) or the CG mixed-organ contrast /χ/-/ħ/ (M=67.2% vs. 64.6%).

To probe that interaction, we looked for individual differences in categorization of Nuu-Chah-Nulth /x/, as we had for English /θ/. In this case, 23 participants formed three subgroups of responders; the 24th split her responses 50/50 between <C> and <H>. The largest subgroup Categorized /x/ above 50% as <H> (n=12; $M_{<H>}$=92.75%), followed by those who Categorized it as Italian <C> above 50% (or in one case as the most frequent choice at 44%) (n=7; $M_{<C>}$=67.14%), and the smallest number Categorized it as <J> (n=4; $M_{<J>}$=70.5%). Thus, the /x/-as-<H> subgroup assimilated Nuu-Chah-Nulth /x/-/χ/ as a CG difference within <H>, but the remaining two subgroups assimilated /x/-/χ/ as a TC-O (overlapping) contrast (<C> or <J> vs. <H>). Therefore, we combined the <C> and <J> categorizers into a single TC-O subgroup (n=11) and conducted an ANOVA on the between-subjects factor Subgroup (<H> vs <C/J> categorizers, i.e., CG vs. TC-O, respectively) and within-subject factors Contrast Type x Consonant that had interacted in the Nuu-Chah-Nulth breakdown analysis. Neither the Subgroup main effect nor any interactions with it were significant. Thus, unlike the case with English /θ/, the Nuu-Chah-Nulth /x/ categorization subgroups did not differ in discrimination performance, not even on the /x/-/χ/ contrast despite their CG vs. TC-O assimilation differences. However, we should note that discrimination of a CG assimilation may not necessarily be expected to be much better that discrimination of a TC-O assimilation type (see Faris, Best, & Tyler, 2016; Fenwick, Best, David, & Tyler, 2017; Tyler, Best, Faber, & Levitt, 2014). Consistent with the Nuu-Chah-Nulth breakdown ANOVA, the significant effects of the current analysis were Contrast Type, $F_{(2, 42)}$=20.884, p<.0001, η_p^2=.499, Consonant (of X in AXB), $F_{(2, 42)}$=4.511, p<.047, η_p^2=.177, and their interaction, $F_{(2, 42)}$=6.272, p<.005, η_p^2=.230.

4. Discussion

The listeners showed Categorized assimilations of each of the English and Nuu-Chah-Nulth consonants to native Italian consonants. Whereas each English fricative was categorized as a different Italian consonant (/f/-as-<F>; /θ/-as-<T>; /s/-as-<S>; /ʃ/-as-<SCIA>), the Nuu-Chah-Nulth fricatives were all Categorized as Italian <H>. Given that Italian (and Venetan Dialect) does not have the phoneme /h/, and that <H> in written Italian words is "silent" [∅], this may mean that the listeners heard *no* consonant at the syllable onsets of the Nuu-Chah-Nulth target stimuli. On the other hand, even though they had been instructed to indicate which *Italian* consonant they perceived, as noted earlier we cannot rule out that they may have chosen <H> to indicate they heard an English [h] for the Nuu-Chah-Nulth consonants given that all but two participants had learned English at school[4]. Those two still chose <H> 94-100% of the time for all Nuu-Chah-Nulth consonants, however, like the L2-English majority. In any case, whether their choices of <H> indicate Italian silent [∅] or English [h], the listeners heard all Nuu-Chah-Nulth fricatives as most similar to the same single category <H>.

The pairwise assimilation patterns were Two Category (TC) for all three English contrasts, two of them non-overlapping (TC-N: /θ/-/s/ and /s/-/ʃ/) and the other overlapping (TC-O: /f/-/θ/), whereas the Nuu-Chah-Nulth contrasts instead showed either Single Category (SC) assimilation (pharyngeal vs. glottal /ħ/-/h/) or Category Goodness difference (CG) assimilation (uvular vs. pharyngeal /χ/-/ħ/, and velar vs. uvular /x/-/χ/). PAM predictions were that the English contrasts should be discriminated significantly better than the Nuu-Chah-Nulth contrasts, which was supported by a main effect of Language. Moreover, both Nuu-Chah-Nulth CG contrasts were predicted by PAM to be discriminated significantly better than the SC contrast, but not to differ from each other, also upheld by the analyses. Thus, overall the PAM predictions were supported quite well.

The core AOH predictions about discrimination levels for within- versus between-organ contrasts, on the other hand, were not supported. Indeed, the observed patterns actually run counter to AOH predictions of better discrimination for between than within-organ non-native contrasts. Performance was better for contrasts involving natively-used and L2-learned articulatory organs, which follows more from PAM principles than AOH predictions. Whereas English-learning infants fail to discriminate

4 One had only learned French, which also lacks [h]; the other had learned no foreign languages.

English anterior fricatives better than posterior Nuu-Chah-Nulth fricatives (Tyler et al., 2014), our Italian adult participants did discriminate the more familiar L2-English fricatives better than the completely unfamiliar non-native Nuu-Chah-Nulth ones.

We must consider, as well, how the results might relate to other models of non-native speech perception. The SLM prediction that *new* phones from a non-native language should be perceived more accurately than *similar* non-native phones, due to the latter being more readily equivalence classified to native phonemes, is contradicted by our finding of significantly *poorer* discrimination for the Nuu-Chah-Nulth fricatives than the English fricatives. However, the present study does not address SLM's predictions that the Nuu-Chah-Nulth fricatives should be more readily established as new L2 phonemes, including more accurate L2 production, as compared to learning and production of English /θ/. Further research would be needed to evaluate those SLM predictions.

Meanwhile, the two English /θ/-assimilation subgroups and their differences in discrimination of English /θ/-/f/ challenge claims that cross-language speech perception is driven primarily by acoustic similarity/distance (e.g., L2LP: Escudero & Boersma, 2005; Holt & Lotto, 2008), given that the listeners who categorized /θ/ to the acoustically more similar /f/ were in the minority rather than the majority, and despite rating the goodness of /θ/ significantly lower than that of /f/ this subgroup showed substantially poorer discrimination of /θ/-/f/ than the majority of listeners who categorized /θ/ to the acoustically more dissimilar Italian /t/. Assimilating /θ/ to Italian /t/ has potential L2 phonological benefits over a more acoustically-based categorization to /f/: it could help L2 Italian learners of English to maintain both the English /θ/-/f/ distinction as a TC assimilation to Italian /t/-/f/, and the English /θ/-/t/ distinction as a CG assimilation between a poor /t/ (2.8 rating for /θ/ as Italian short-lag /t/) versus a moderately good /t/ (English long-lag /t/ as Italian /t/).

Further studies comparing perception of these English versus Nuu-Chah-nulth contrasts by listeners of varying L1s that differ in contrastive use of the tongue body, tongue root, epiglottis and glottis for voiceless fricatives could further delineate the contributions of perceptual assimilation (PAM) and equivalence classification (SLM) on categorization and discrimination of non-native and L2 consonant contrasts. Studies on the impact of L2 learning or laboratory perceptual training on perception and production of the posterior fricative series by different L1 groups would also be informative.

5. References

Abramson, A. S., & Lisker, L. (1970). Discriminability along the voicing continuum: Cross-language tests. *6th International Congress of Phonetic Sciences, Prague*, 569-573.

Abramson, A. S., & Lisker, L. (1973). Voice-timing perception in Spanish word-initial stops. *Journal of Phonetics*, *1*, 1-8.

Antoniou, M., Best, C. T., & Tyler, M. D. (2013). Greek-English bilinguals' and Greek and English monolinguals' perception of nonnative Ma'di voicing contrasts. *Journal of the Acoustical Society of America*, *133,* 2397-2411.

Antoniou, M., Tyler, M. D., & Best, C. T. (2012). Two ways to listen: Do bilinguals perceive stop voicing differently according to language mode? *Journal of Phonetics*, *40*, 582-594.

Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2010). Language context elicits native-like stop voicing in early bilinguals' productions in both L1 and L2. *Journal of Phonetics*, *38*, 640-653.

Antoniou, M., Best, C. T., Tyler, M. D., & Kroos, C. (2011). Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics, 39*, 558-570.

Avesani, C., Galatà, V., Vayra, M., Best, C. T., Di Biase, B., Tordini, O., & Tisato, G. (2016). In M. Vayra, C. Avesani, & F. Tamburini (Eds.), *Italian roots in Australian soil: Coronal obstruents in native dialect speech of Italian-Australians from two areas of Veneto* [*Language acquisition and language loss: Acquisition, change and disorders of the language sound structure*] (pp. 74-98). Milan: Studi AISV. ISBN: 978-88-97657-11-8.

Avesani, C., Galatà, V., Best, C. T., Vayra, M., Di Biase, B., & Ardolino, F. (2017). Phonetic details of coronal consonants in the Italian spoken by Italian-Australians from two areas of Veneto. In C. Bertini, C. Celata, G. Lenoci, C. Meluzzi, & I. Ricci (Eds.), *Social and biological factors in speech variation* (pp. 281-306). Milan, Studi AISV. ISBN 978-88-97657-19-4.

Bertinetto, P. M., & Loporcaro, M. (2005). The sound pattern of Standard Italian, as compared with the varieties spoken in Florence, Milan and Rome. *Journal of the International Phonetics Association*, *35*(2)*,* 131-151.

Best, C. T. (1995). A direct realist perspective on cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 167-200). Timonium, MD, York Press.

Best, C. T. (2019). The diversity of tone languages and the roles of pitch variation in non-tone languages: Considerations for tone perception research. *Frontiers in Psychology*. (online publication).

Best, C. T., Goldstein, L., Tyler, M. D., & Nam, H. (2016). Articulating what infants attune to in native speech. *Ecological Psychology, 28*, 216-261.

Best, C. T., & McRoberts, G. W. (2003). Infant perception of nonnative consonant contrasts that adults assimilate in different ways. *Language & Speech*, *46,* 183-216.

Best, C. T., McRoberts, G. W., & Goodell, E. (2001). American listeners' perception of nonnative consonant contrasts varying in perceptual assimilation to English phonology. *Journal of the Acoustical Society of America*, *1097*, 775-794.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 45-60.

Best, C. T., Shaw, J., & Clancy, E. (2013). Recognizing words across regional accents: The role of perceptual assimilation in lexical competition. *Interspeech*, *2013*, 2128-2132.

Best, C. T., Shaw, J., Mulak, K., Docherty, G., Evans, B., Foulkes, P., Hay, J., Al-Tamimi, J., Mair, K., & Wood, S. (2015a). Perceiving and adapting to regional accent differences among vowel subsystems. *18th International Congress of Phonetic Sciences, Glasgow*, 0964. http://www.icphs2015.info/pdfs/Papers/ICPHS0964.pdf

Best, C. T., Shaw, J., Mulak, K., Docherty, G., Evans, B., Foulkes, P., Hay, J., Al-Tamimi, J., Mair, K., & Wood, S. (2015b). From Newcastle MOUTH to Aussie ears: Australians' perceptual assimilation and adaptation for Newcastle UK vowels. *Interspeech, 2015,* 1932-1936.

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

Best, C. T., Traill, A., Carter, A., Harrison, K. D., & Faber, A. (2003). !Xóõ click perception by English, Isizulu, and Sesotho listeners. *15th International Congress of Phonetic Sciences, Barcelona*, 853-856.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In M. Munro, & O.-S. Bohn (Eds.), *Second Language Speech Learning* (pp. 13-34). Amsterdam: John Benjamins Publishing.

Boersma, P., & Weenink, D. (2009). *Praat: Doing phonetics by computer* [Computer program]. Retrieved 2009. http://www.fon.hum.uva.nl/praat/

Bohn, O.-S., & Flege, J. E. (1990). Interlingual identification and the role of foreign language experience in L2 vowel perception. *Applied Psycholinguistics*, *11*, 303-328.

Bohn, O.-S., & Flege, J. (1993). Perceptual switching in Spanish/English bilinguals: Evidence for universal factors in stop voicing judgments. *Journal of Phonetics*, *21*, 267-290.

Bundgaard-Nielsen, R. L., Best, C. T., & Tyler, M. D. (2011a). Vocabulary size is associated with second language vowel perception performance in adult second language learners. *Studies in Second Language Acquisition*, *33*, 433-461.

Bundgaard-Nielsen, R. L., Best, C. T., Kroos, C., & Tyler, M. D. (2011b). Vocabulary size matters: The assimilation of L2 Australian English vowels to L1 Japanese vowel categories. *Applied Psycholinguistics*, *32*, 51-67.

Bundgaard-Nielsen, R. L., Best, C. T., Kroos, C., & Tyler, M. D. (2012). Second language learners' vocabulary expansion is associated with improved Second Language vowel intelligibility. *Applied Psycholinguistics*, *33*, 643-664.

Carlson, B. F., & Esling, J. H. (2003). Phonetics and physiology of the historical shift of uvulars to pharyngeals in Nuuchahnulth (Nootka). *Journal of the International Phonetic Association*, *33*, 183-193.

Curtin, S., Byers-Heinlein, K., & Werker, J. F. (2011). Bilingual beginnings as a lens for theory development: PRIMIR in focus. *Journal of Phonetics*, *39*, 492-504.

Dupoux, E., & Peperkamp, S. (2002). Fossil markers of language development: Phonological 'deafnesses' in adult speech processing. In J. Durand, & B. Laks (Eds.), *Phonetics, phonology, and cognition* (pp. 168-190). Oxford, UK: Oxford University Press.

Escudero, P., & Boersma, P. (2004). Bridging the gap between L2 speech perception research and phonological theory. *Studies in Second Language Acquisition*, *26*, 551-585.

Faris, M. M., Best, C. T., & Tyler, M. D. (2018). Discrimination of uncategorised non-native vowel contrasts is modulated by perceived overlap with native phonological categories. *Journal of Phonetics*, *70*, 1-19.

Faris, M. M., Best, C. T., & Tyler, M. D. (2016). An examination of the different ways that non-native phones may be perceptually assimilated as uncategorized. *JASA-Express Letters*, *139*, EL1-5.

Fenwick, S. E., Best, C. T., Davis, C., & Tyler, M. D. (2017). The influence of auditory-visual speech and clear speech on cross-language perceptual assimilation. *Speech Communication*, *92*, 114-124. http://dx.doi.org/10.1016/j.specom.2017.06.001

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 233-277). Timonium, MD: York Press.

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. (2007). Language contact in bilingualism: Phonetic system interactions. In J. Cole, & J. Hualde (Eds.), *Laboratory Phonology 9* (pp. 353-380). Berlin: Mouton de Gruyter.

Goldstein, L. M., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. In N. O. Schiller, & A. S. Meyer (Eds.), *Phonetics and phonology in language comprehension and production: Differences and similarities* (pp. 159-207). Berlin: Mouton de Gruyter.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *The Journal of the Acoustical Society of America*, *107*, 2711-2724.

Hallé, P. A., Chang Y.-C., & Best, C. T. (2004). Identification and discrimination of Mandarin Chinese tones by Chinese versus French listeners. *Journal of Phonetics*, *3*, 395-421.

Hammond, R. M. (2001). *The sounds of Spanish: Analysis and application*. Somerville, MA: Cascadilla.

Holt, L. L., & Lotto, A. J. (2008). Speech perception within an auditory cognitive science framework. *Current Directions in Psychological Science*, *17*, 42-46.

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

Jusczyk, P. W. (1993). From general to language-specific capacities: The WRAPSA model of how speech perception develops. *Journal of Phonetics*, *21*, 3-28.

Jusczyk, P. W. (1997). *The discovery of spoken language.* Cambridge, MA: MIT Press.

Krebs-Lazendic, L., & Best, C. T. (2013). First language suprasegmentally-conditioned syllable length distinctions influence perception and production of second language vowel contrasts. *Laboratory Phonology*, *4*, 435-474.

Kuhl, P. K. (1993a). Innate predispositions and the effects of experience in speech perception: The native language magnet theory. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. MacNeilage, & J. Morton (Eds.), *Developmental neurocognition: Speech and face processing in the first year of life* (pp. 259-274). Dordrecht, the Netherlands: Springer.

Kuhl, P. K. (1993b). Early linguistic experience and phonetic perception: Implications for theories of developmental speech perception. *Journal of Phonetics*, *21*, 125-139.

Kuhl, P. K., Conboy, B. T., Coffey-Corina, S., Padden, D., Rivera-Gaxiola, M., & Nelson, T. (2008). Phonetic learning as a pathway to language: New data and native language magnet theory expanded (NLM-e). *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *363*, 979-1000.

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*, F13-F21.

Lecumberri, M. L. G., Cooke, M., & Cutler, A. (2010). Non-native speech perception in adverse conditions: A review. *Speech Communication*, *52*, 864-886.

Leussen, V., & Escudero, P. (2015). Learning to perceive and recognize a second language: The L2LP model revised. *Frontiers in Psychology*, *6*, 1000.

MacKain, K. S., Best, C. T., & Strange, W. (1981). Categorical perception of /r/ and /l/ by Japanese bilinguals. *Applied Psycholinguistics*, *2*, 369-390.

Mioni, A. (2001). *Elementi di fonetica*. Padova: Unipress.

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

Polivanov, E. (1931/1974). La perception des sons d'une langue étrangère. *Travaux du Cercle linguistique de Prague*, *4*, 79-96. Translated by D. Armstrong (1974). *Polivanov E. Selected Works. Articles on general linguistics.* Paris: Mouton.

Polka, L. (1991). Cross-language speech perception in adults: Phonemic, phonetic, and acoustic contributions. *Journal of the Acoustical Society of America*, *89*, 2961-2977.

Polka, L. (1992). Characterizing the influence of native language experience on adult speech perception. *Perception & Psychophysics*, *52*, 37-52.

Polka, L., & Bohn, O.-S. (2003). Asymmetries in vowel perception. *Speech Communication*, *41*, 221-231.

Polka, L., & Bohn, O.-S. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, *39*, 467-478.

Polka, L., Colantonio, C., & Sundara, M. (2001). A cross-language comparison of/d/–/ð/perception: evidence for a new developmental pattern. *The Journal of the Acoustical Society of America*, *109*, 2190-2201.

Pollack, I., & Pisoni, D.B. (1971). On the comparison between identification and discrimination tests in speech perception. *Psychonomic Science*, *24*, 299-300.

Reid, A., Burnham, D., Attina, V., Kasisopa, B., Schwarz, I.-C., & Best, C. T. (2015). Perceptual assimilation of lexical tone: The role of language experience and visual information. *Attention, Perception & Psychophysics*, *77*, 571-591.

Sebastián-Gallés, N., & Soto-Faraco, S. (1999). Online processing of native and non-native phonemic contrasts in early bilinguals. *Cognition*, *72*, 111-123.

Shaw, J. A., Best, C. T., Docherty, G., Evans, B., Foulkes, P., & Hay, J. (2018). Resilience of English vowel perception across regional accent variation. *Laboratory Phonology*, *9*(1), 11, 1–36. http://doi.org/10.5334/labphon.87.

Shaw, J., Best, C. T., Mulak, K., Docherty, G., Evans, B., Foulkes, P. Hay, J., Al-Tamimi, J., Mair, K., Peek, M., & Wood, S. (2014). Effects of short-term exposure to unfamiliar regional accents: Australians' categorization of London and Yorkshire English consonants. *Speech Science and Technology*, *2014*, 71-74.

So, C., & Best, C. T. (2010). Cross-language perception of non-native tonal contrasts: Effects of native phonological and phonetic influences. *Language & Speech*, *53*, 273-293.

So, C., & Best, C. T. (2011). Categorizing Mandarin tones into listeners' native prosodic categories: The role of phonetic properties, *Poznañ Studies in Contemporary Linguistics*, *47*, 133-145.

So, C., & Best, C. T. (2014). Phonetic influences on English and French listeners' assimilation of Mandarin tones to native prosodic categories. *Studies in Second Language Acquisition*, *36*, 195-221.

Strange, W. (1995). Cross-language studies of speech perception: A historical review. In W. Strange (Ed.), *Speech perception and linguistic experience: Theoretical and methodological issues in cross-language speech research* (pp. 3-45). Timonium, MD: York Press.

Strange, W. (2011). Automatic selective perception (ASP) of first and second language speech: A working model. *Journal of Phonetics*, *39*, 456-466.

Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: The re-education of selective perception. In J. G. H. Edwards, & M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 153-192). Amsterdam: John Benjamins.

Studdert-Kennedy, M., & Goldstein, L. (2003). Launching language: The gestural origin of discrete infinity. In M. H. Christensen, & S. Kirby (Eds.), *Language Evolution* (pp. 235-254). Oxford, UK: Oxford University Press.

Trubetzkoy, N. S. (1939/1969). Grundzüge der phonologie, *Travaux du Cercle Linguistique de Prague*, *7*, 1–271. Translated by C.A.M. Baltaxe (1969). *Principles of phonology.* Berkeley, CA: University of California Press.

Tsao, F. M., Liu, H. M., & Kuhl, P. K. (2006). Perception of native and non-native affricate-fricative contrasts: Cross-language tests on adults and infants. *Journal of the Acoustical Society of America*, *120*, 2285-2294.

Tyler, M. D., Best, C. T., Faber, A., & Levitt, A. G. (2014). Perceptual assimilation and discrimination of non-native vowel contrasts. *Phonetica*, *71*, 4-21.

Tyler, M. D., Best, C. T., Goldstein, L. M., & Antoniou, M. (2014). Investigating the role of articulatory organs and perceptual assimilation in infants' discrimination of native and non-native fricative place contrasts. *Developmental Psychobiology*, *56*, 210-227.

Werker, J. F., & Curtin, S. (2005). PRIMIR: A developmental framework of infant speech processing. *Language learning and development*, *1*, 197-234.

Werker, J. F., Gilbert, J. H., Humphrey, K., & Tees, R. C. (1981). Developmental aspects of cross-language speech perception. *Child Development*, *52*, 349-355.

Zamboni, A. (1974). *Veneto*. Pisa: Pacini Editore.

Zamboni, A. (1988). Italienisch: Areallinguistik IV a) Venezien. In G. Holtus, M. Metzeltin, & C. Schmitt (Eds.), *Lexicon der Romanistischen Linguistik*, vol. IV (pp. 517-538). Tubingen, Germany: Niemeyer Verlag.