Article:

Dar, Mariam, Keren-Portnoy, Tamar orcid.org/0000-0002-7258-2404 and Vihman, Marilyn orcid.org/0000-0001-8912-4840 (2018) An order effect in English infants' discrimination of an Urdu affricate contrast. Journal of Phonetics. pp. 49-64. ISSN 0095-4470

https://doi.org/10.1016/j.wocn.2017.12.002

An order effect in English infants' discrimination of an Urdu affricate contrast

Mariam Dar*, Tamar Keren-Portnoy and Marilyn Vihman

*University of York, York YO10 5DD,
Corresponding author: md738@york.ac.uk, Tel: +447446905025

Abstract
An order effect was found in English infants' discrimination of an Urdu contrast. In Experiment 1 7- and 11-month-old English infants were tested on the Urdu contrast between the affricates /tʃʰ/ and /tʃ/. The order of presentation was counterbalanced: At each age half the infants were habituated to the aspirated and tested on the unaspirated affricate, the other half habituated to the unaspirated and tested on the aspirated. As expected, younger infants discriminated the contrast whereas older infants did not, showing the expected decline in discrimination. Order of presentation seemed to affect the older infants' response. Experiment 2 tested the order effect directly. The results showed no asymmetry in the performance of 7-month olds but clear asymmetry in that of 11-month-olds, who discriminated the contrast only when the non-English-like aspirated affricate was presented first. Experiment 3 tested adult native-speakers of both Urdu and English. Although the English listeners showed a reduced sensitivity to the contrast, there was no effect due to order of presentation of the stimuli in either adult group. The finding of an asymmetry in the infants suggests that infants' perceptual narrowing for speech sounds may be a more complex phenomenon than has generally been assumed.

Keywords
Urdu voiceless affricate contrast, order effect in consonant discrimination, perceptual narrowing

An order effect in English infants' discrimination of an Urdu affricate contrast

1.0 Introduction

There is ample evidence that infants are born with 'universal' listening abilities that allow them to successfully discriminate most of the phonetic contrasts found in the world's languages. This ability is not maintained into adulthood, however; a developmental decline in the discrimination of contrasts has been demonstrated in numerous studies, using a range of non-native contrasts (Best & McRoberts, 2003; Best, McRoberts, LaFleur & Silver-Isenstadt, 1995; Bohn & Polka, 2001; Kuhl, Conboy, Coffey-Corina, Padden, Rivera-Gaxiola & Nelson, 2008; Kuhl, Stevens, Hayashi, Deguchi, Kiritani & Iverson, 2006; Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992; Werker & Tees, 1983, 1984; Werker & Tees, 2002; Werker, Gilbert, Humphrey, & Tees, 1981).[1] Ambient language exposure enables infants to form phonetic and phonological categories that affect the way they perceive the sounds around them (Maye, Weiss, & Aslin, 2008; Maye, Werker, & Gerken, 2002; Wanrooij, Boersma & van Zuijen, 2014). Interestingly, studies have also shown that despite a significant shift in perception before the first birthday, the ability to perceive non-native contrasts is not entirely lost, and only minimal exposure is required to reinstate sensitivity during the period of decline (Conboy, Sommerville & Kuhl, 2008; Kuhl, Tsao & Liu, 2003; Maye et al., 2008; Yeung & Werker, 2009; Yoshida, Pons, Maye & Werker, 2010). The current study investigates this decline in discrimination for a non-native consonant contrast. In particular, it reveals the fact that the experimentally robust finding of such a decline in infant discrimination is, at least in part, dependent on task characteristics, and specifically, on the order in which sounds are presented to the infant in the experimental task.

Werker et al. (1981) were the first to attempt to trace the time-course of perceptual decline from infancy to adulthood. English- and Hindi-speaking adults and English-learning infants were tested on two Hindi contrasts: (1) dental /ta/ vs. retroflex /ʈa/, and (2) voiceless aspirated /tʰ/ vs. breathy-voice /dʰ/. Prior to the experiment ten English-speaking adults had received training in the Hindi contrasts. Hindi adults and English-learning infants were able to discriminate the sounds, but only one English-speaking adult could perceive the difference without prior training. This study also pointed towards the fact that in adults there is a decline rather than a complete loss of the ability to discriminate non-native contrasts, since discrimination remained possible with specific training. Werker and Tees (1983) further tested English-speaking children at four, eight and twelve years of age: None of the children were able to discriminate the Hindi contrasts. In a subsequent series of experiments with infants within the first year of life (Werker & Tees, 1984) it was found that most 6- and 8-10-month-olds could discriminate the non-native contrast from Hindi and a new pair of unfamiliar consonants from Thompson (or Nthlakampx), while most 10- to 12-month-olds could not. Werker and Tees concluded that a 'selective tuning of initial sensitivities in accordance with a

[1] Note that not all contrasts show this trajectory: Some sounds (notably, fricatives) are not well discriminated in early infancy (Aslin & Pisoni, 1980; but see Tsao et al., 2006), others are discriminated at an early age but show later enhancement in the ability to discriminate them (e.g., /r/ vs. /l/, see Kuhl et al., 2006). Aslin and Pisoni (1980) describe several possible developmental consequences of the interaction between infants' perceptual abilities at birth and the language input that they receive. We limit ourselves here to discussing cases in which a decline is seen.

specific phonology...occurs at about the age that the child is beginning to understand and possibly produce sounds appropriate to his/her native language' (1984, p. 62).

This reorganization, currently referred to as 'perceptual narrowing' (Kuhl, 2004), leads to a decline in the discrimination of non-native consonant contrasts. As indicated above, parallel results have been obtained in a variety of behavioral studies using non-native phonetic consonant contrasts from languages such as Hindi, Japanese, Mandarin, Spanish, Swedish and Zulu (for a summary, see Werker & Tees, 2002) as well as in neural imaging studies (using Event Related Potentials: Rivera-Gaxiola, Klarman, Garcia-Sierra & Kuhl, 2005a, Rivera-Gaxiola, Silva-Pereyra & Kuhl, 2005b), which have yielded similar results. The age at which the decline occurs, toward the end of the first year, is consistent across all of these studies. The decline for non-native vowels occurs much earlier, according to some studies. For example, Polka and Werker (1994) showed a decline for English infants' discrimination of non-native German vowels between 6-8 months of age (see also Bosch and Sebastián-Gallés, 2003; Tsuji and Cristia, 2014), but other studies paint a more nuanced picture, with no decline at all by age 12 months (Polka & Bohn, 1996), or a decline followed by a recovery of the ability to detect the contrast by age 12 months (in bilingual infants: Bosch and Sebastián-Gallés, 2003).

In this study, however, we will focus on findings regarding the decline in discrimination for consonants, which has aroused less controversy. This decline can be understood to be the result of early language learning experience leading to category learning. Taking an exemplar model perspective on language learning, we can think of early perceptual abilities as engaging a parametric phonetic level of perception (see Pierrehumbert, 2003), which is graded, continuous, multidimensional and multimodal and which does not as yet involve any categories. Infants listening in this way are able to distinguish between sounds that are different in their auditory properties, even if the differences are small. At this early stage, infants can be seen as 'universal' listeners, not because they somehow anticipate (or 'possess' or 'know') all the possible sound contrasts, but because they do not listen for contrast or categories at all. Further language exposure will teach the infants which differences they can ignore or forget – namely, differences that play no functional or phonemic role in the native language.

Putting it simply for the sake of exposition, we can assume that if two sounds contrast in a language (i.e., if they are distinct phonemes), their phonetic realization will be bimodally distributed (Maye et al., 2008; Pierrehumbert, 2003). This leads infants to form separate phonetic categories, one for each of the sounds. When the speech sounds do not contrast, input speech is likely to provide a broader or unimodal range of variation. Various studies have provided evidence of such learning. For example, Maye et al. (2002) presented 6- and 8-month old infants with a continuum from prevoiced to voiceless unaspirated alveolar stops; one group received unimodal exposure, the other group bimodal exposure. At both ages, only the infants given bimodal exposure discriminated the stops successfully when tested later, showing that two separate phonetic categories were formed only in that condition.[2]

[2] It should be noted, however, that there is no information as to the way the infants perceived the continuum prior to their exposure to either of the conditions. This makes the interpretation of category formation in this study inconclusive. We thank an anonymous reviewer for pointing this out.

Category learning can influence the way speakers discriminate both native and non-native language contrasts. We take the representations of sounds to be built from multiple exemplars organized in a multidimensional network, with denser and sparser areas, or with different exemplars and different areas more or less strongly connected to others. The dense areas, or densely connected clusters, are categories. The sparser ones are either category peripheries or lie outside of any category (see Iverson and Kuhl, 1995). In such a framework, we can expect the categorizations of sounds to show effects of prototypicality, centrality, and fuzzy boundaries. And such effects have indeed been found for vowels. Vowel categories are understood to be built around prototypes (Grieser & Kuhl, 1989) and to have graded membership and fuzzy boundaries (Taylor, 2008). Kuhl and her colleagues (Grieser & Kuhl, 1989; Kuhl 1986, 1991) have proposed that discrimination is affected by the relationship between the 'most typical' (prototypical) tokens of a given category and the less typical ones. That is, for vowel categories some areas in the perceptual space serve as 'category centers', providing a reference point for generalization to novel exemplars. In a test of adult perception of differences between the within-category exemplars of the vowel /i/ adults rated different variants of the stimuli on 'goodness' (defined as sounding 'natural'). The variants near the center received significantly higher rankings than the variants near the category boundary. The same stimuli were then used to test two groups of 6-month-olds. One group was tested on discrimination between the 'good' /i/ and its variants, based on the adult judgment scores, and the other on the 'poor' /i/ and its variants. Infants discriminated significantly more often between the non-prototypical than between the prototypical exemplars. Thus, the stimuli that adults ranked as good exemplars of the category resulted in greater infant generalization to other members of the same vowel category than the stimuli that adults ranked as poor. Kuhl and colleagues interpreted these results as suggesting that the prototype serves as a perceptual 'magnet' for the sound category; that is, it 'pulls' other stimuli towards it, effectively shortening the perceptual distance between them. (In its latest version, this model is called the Native Language Magnet theory expanded, or NLM-e: Kuhl et al., 2008.)

Similar to Kuhl's model, the Perceptual Assimilation Model (PAM: Best, 1993) describes a process by which listeners perceptually assimilate non-native sounds (whether vowels or consonants) to their own phonemic inventory, based on their experience with the native language; non-native sounds are categorized or identified in relation to where they fall within the network of native language sounds. According to this model, whether infants do or do not distinguish between non-native sounds has to do with whether those sounds fall near a native category prototype, near its periphery, between categories, or completely outside the area of what counts as language sounds (as in the case of clicks, which are perceived as 'non-assimilable' by listeners in whose native language they lack segmental status). In the case of PAM, the dimensions according to which similarity or distance is computed are articulatory. (Best, 1993, assumes that what listeners perceive in listening to speech sounds are the articulatory gestures that underlie their production.) Given that the parametric space we described earlier is multimodal, this does not change the developmental picture we are describing here but only enriches it.

Another finding that can be explained within the framework of exemplar models is that of order-related asymmetry in discrimination between more prototypical and less prototypical tokens within a vowel category (such asymmetries have been reported for native language vowel contrasts for adult speakers: Cowan and Morse, 1986; Iverson & Kuhl, 1995; Repp, Healy, & Crowder, 1979). In a study of vowel discrimination in infants (Polka & Werker, 1994) an asymmetry was found in which a given direction of change ([y] preceding [u] and [ʏ] preceding [ʊ]) was discriminated better by English infants than the reverse order. That is, 6-8-month-old infants discriminated the vowels only when the 'non-prototypical' (non-English-like) front-rounded vowel was presented before the 'prototypical' (English-like) back-rounded vowel. The authors first attributed the results to the magnet effect: [u] and [ʊ] are the more familiar vowels for English listeners, while [y] and [ʏ] are non-native; thus, as an anchor point [u] pulled in the perception of the following [y], resulting in assimilation when it was presented first.

However, Polka and Bohn (1996) later suggested that the anchor point plays a role independently of the status of a given vowel in native language phonology. In their study 6-8 and 10-12-month-olds from English and German families were tested on an English (non-German) /ɛ/-/æ/ contrast and a German (non-English) /u/-/y/ contrast. Discrimination was found to be easier from /y/ to /u/ and /ɛ/ to /æ/ than from /u/ to /y/ and /æ/ to /ɛ/, with age or native language not affecting the results. In another study (Polka and Bohn, 2011) large numbers of Danish-learning infants of 6-9 months of age were tested on a Southern British-English contrast, peripheral /ɑ/ vs. /ʌ/, and two native contrasts, /e/ vs. /ɛ/ and /e/ vs. /ø/. It was found that both younger and older children discriminated the non-native vowel more successfully when the less peripheral/more central vowel was presented first. The authors attributed the asymmetry to an innate perceptual bias not dependent on language experience or familiarity from the native language. The findings led Polka and Bohn (2011) to introduce the Natural Referent Vowel (NRV) model, which speculates that 'vowels with extreme articulatory-acoustic properties…act as natural referent vowels…by attracting infant attention and providing stable perceptual forms' (p. 474).

It seems, therefore, that two types of asymmetries in segment discrimination have been shown to exist: one that is likely to be an outcome of biases inherent to the perceptual system (or to its interactions with the production system, as postulated by Stevens - e.g., Stevens & Keyser, 2010), and another that is the result of learning, following exposure to a particular ambient language. The first type of asymmetry should be evidenced from the earliest ages, as it is thought to be a product of the physical/perceptual mechanisms available to infants. The second type is likely to develop with age as a result of increased exposure to a given language. It is this second type of asymmetry that this project is focused on.

Note that asymmetries in vowel perception have been investigated, based on the understanding that vowels are not rigidly organized into clear-cut categories (see e.g., Pierrehumbert, 2003). Consonants, however, have traditionally been assumed to have all-or-none membership or better-defined boundaries, a view which fits with the finding of infant categorical perception (Eimas, Siqueland, Jusczyk & Vigorito, 1971; see also Damper & Harnad, 2000; Livingston, Andrews & Harnad, 1998). Despite the view that consonants are perceived categorically, a number of studies have not only reported that within-category tokens may be discriminable (Miller, 1994) but also that these within-category distinctions affect lexical processes (Dahan, Magnuson, Tanenhaus & Hogan, 2001; McMurray, Tanenhaus & Aslin, 2002). This supports a view of consonant categories as built from exemplars, with the concomitant implication that the categories are centered around prototypes, with graded membership and fuzzy boundaries (see also Pierrehumbert, 2003).

McMurray and Aslin (2005) familiarized 8-month-olds with one member of each of several minimal pairs (e.g., *pear – bear*). Half the infants were familiarized with a word with a voiced onset stop and half with a word with an unvoiced onset stop. Infants were then tested on those same words, their minimal pair and a variant of the familiarized word whose onset, though still within the same voicing category, was shifted towards the other voicing category. Infant looking times showed that they distinguished not only between the between-category variants but also between the within-category variants. A sizeable body of work with adults has also provided evidence against the strong version of categorical perception (Carney, Widin, & Viemeister, 1977; Miller, 1997; Pisoni & Lazarus, 1974; Pisoni & Tash, 1974). Indeed, both Kuhl and colleagues' NLM-e model and Best's PAM refer explicitly to consonants as well as to vowels, treating them too as being organized into the same kinds of categories as vowels, centered around prototypes, with fuzzy boundaries. Taken together in the context of consonants, both models speculate that non-native consonant sounds are perceived in relation to their similarity to consonants in the native language inventory. The studies demonstrating the decline in discrimination for non-native contrasts have shown that any non-native consonant that is similar, if not identical, to a consonant in the ambient language will be perceptually assimilated to that close native consonant sound. In the light of NLM, the native language prototype may be acting as a magnet in these cases, pulling the perception of the non-native consonant towards itself, shortening the perceptual distance and leading to poor discrimination. Moreover, even when a non-native consonant pair is perceived as belonging to the single native category, one consonant will be perceived as a better fit than the other (Best, 1993). If we accept, then, that infant perception is not as 'categorical' as has previously been suggested, could the asymmetries found in vowel perception be found in the case of consonant perception as well? That is, do infants show within-category order effects that might reflect prototypicality? Would infants fail to discriminate a prototypical exemplar from a subsequently presented less prototypical exemplar, but be able to discriminate the same two exemplars when they are presented in the reverse order? Note that there have been reports of asymmetry in discrimination between two consonants that are phonemically distinct in a child's language (e.g., Altvater-Mackensen & Fikkert, 2010; Tsuji et al., 2015; Nam & Polka, 2016). This kind of asymmetry within the native language is not what concerns us here; we aim to test possible asymmetry in non-native consonant perception, i.e., in perception of two segments which may have no straight-forward mapping onto distinct native language categories.

Two studies have reported asymmetries in non-native consonant perception, but with conflicting results. Kuhl et al. (2006) reported asymmetries for 6-8- and 10-12-month-old American and Japanese infants in response to /la/ – /ra/ stimuli. The study found a directional asymmetry *regardless of age or language experience*: Infants found it easier to detect a stimulus change from /la/ to /ra/ than the reverse. This type of finding is perhaps best explained by models which refer to a universal perceptual bias, such as Polka and Bohn's Natural Referent Vowel (NRV) model (2011). In contrast, in Segal, Hejli-Assi, and Kishon-Rabin (2016) consonant asymmetry was found to be dependent on *both age and native language*. Segal and her colleagues tested the discrimination of the voicing contrast /ba/-/pa/ in Arabic-learning infants (whose native language has /b/ but not /p/) and Hebrew-learning infants (whose native language includes a phonological contrast between /p/ and /b/) at 4-6 and 10-12 months of age. The Hebrew-learning infants discriminated the contrast at both ages; no directional asymmetry was observed. On the other hand, there was a decrease in perception of the non-native contrast by the Arabic-learning children between 4-6 and 10-12 months of age. In addition, at 10-12 months of age Arabic-learning infants failed to discriminate the change from /ba/ to /pa/ but showed a marginally significant effect for the change from /pa/ to /ba/; no such asymmetries were found at 4-6 months of age. Though the effect in that study was only marginal, its direction was consistent with the predictions of the PAM model and to those of the NLM-e model in relation to vowels: For the Arabic-learning infants, the /pa/ tokens could have been perceived as atypical examples of /ba/, whereas the /ba/ tokens were prototypical exemplars. As a result, when the atypical /pa/ was presented first, the infants discriminated between the two syllable types, but when the order was reversed they did not.

The present study was initially designed to explore developmental change in both native and non-native consonant perception in infants at the end of their first year, as part of a longitudinal study of English-learning and Urdu-learning infants. In the course of the study we discovered some unexpected but intriguing signs of asymmetry in the perception of consonants. We therefore pursued our investigation of the issue. This paper will report only on the study with English-learning infants and adults tested on a non-native contrast from Urdu, [3] /tʃ/ vs. /tʃʰ/. Experiment 1 was conducted to determine whether there was a decline in English infants' perception of non-native Urdu affricate contrast between 7-11 months of age, as reported in the literature for other non-native consonant contrasts. Experiment 2 tested for order effects in English infants' discrimination of the non-native contrast, after a trend was observed in Experiment 1. Lastly, Experiment 3 was conducted on English and Urdu adults to test whether the order effects were maintained after infancy. Note that affricates and fricatives have been used in fewer studies as compared to other consonantal contrasts (Eilers & Minife, 1975; Levitt et al., 1987; Tsao, Liu, Kuhl, & Tseng, 2000; Polka, Colantonio, & Sundara, 2001; Ting et al., 2006; Johnson & Babel, 2007; Beach et al., 2008), perhaps because several early studies conducted with young infants failed to provide evidence of discrimination of fricatives (Vihman, 1996). Tsao et al. (2006) is an exception; this investigation tested Chinese and English infants on a Mandarin affricate-fricative contrast /tɕʰ/ vs. /ɕ/ and showed discrimination at 6-8 months in both groups. However, no study to date has tested English infants on a contrasting affricate pair.

2.0 Experimental materials and methods – Pilot

2.1 Pilot study: Materials and methods

2.1.1 Choosing an Urdu contrast for testing with English listeners

Adult English speakers were tested on Urdu minimal pairs differing in the feature of aspiration (e.g., /bʰʌr/ 'to fill' and /bʌr/ 'groom') and on singleton vs. geminate consonants, to establish which contrast was the most difficult to discriminate. The contrast with the largest number of incorrect responses and the longest response times was then tested on 7- and 11-month-olds from English-speaking homes.

2.1.2 Participants

[3] Urdu and Hindi, despite being spoken in two different regions and having different names and orthography, are essentially the same language, with minor differences in lexicon but similar phonology.

Twenty adult monolingual English speakers, most of them university students (age range 22-26 years), were recruited for the experiment.

2.1.3 Stimuli

The stimuli included 11 Urdu consonant contrasts (differing by the presence or absence of aspiration) and 5 geminate-singleton contrasts; none of these contrasts exists in English (see Table 1). A female native Urdu speaker recorded the stimuli. We created 25 ‘different’ word pairs (minimal pairs) and 20 ‘same’ pairs, representing 16 phonemic contrasts of Urdu. Each of the words was recorded with a carrier sentence ‘can you say’ (تم یہ بولو). Each word was recorded three times to provide three different tokens of each word.

Table 1 List of Urdu aspirate/non-aspirate and geminate/singleton contrasts used in the pilot

Place of articulation	*Aspirate/non-aspirate (initial only)*	*Number of minimal pairs*	*Singleton/geminate (medial only)*	*Number of minimal pairs*
Bilabial	/p/ - /p^h/	2		
	/b/ - /b^h/	4		
Dental	/t/ - /t^h/	3	/t/ - /t:/	1
	/d/ - /d^h/	2	/d/ - /d:/	1
Post-alveolar	/tʃ/ - /tʃ^h/	2	/tʃ/ - /tʃ:/	1
	/dʒ/ - /dʒ^h/	2		
Retroflex	/ʈ/ - /ʈ^h/	2	/ʈ/ - /ʈ:/	1
	/ɖ/ - /ɖ^h/	2		
	/ɽ/ - /ɽ^h/	2		
Velar	/k/ - /k^h/	2	/k/ - /k:/	1
	/g/ - /g^h/	2		

2.1.4 Method

The adult participants were tested with an AX discrimination task using E-Prime. Each participant was auditorily presented with 44 pairs of Urdu words over sound-cancelling Bose QC-15 headphones and asked to judge whether they were the same or different, beginning with three practice trials. Participants were asked to press a key (‘s’) to indicate ‘same’, if the two sounds in a pair seemed to be identical, and another key (‘d’) if the sounds were judged to be different. The inter-stimulus interval between contrasts was one second, the intra-stimulus gap 300 milliseconds. The order of stimuli was randomized. Each word pair was presented once in each of four combinations: AB (word A followed by word B), BA, AA and BB, with no recorded token of any word being used more than once. In the practice trials, the participant was taken to the next pair only when the correct key had been pressed. The test trials then started automatically. As soon as the participants pressed a response key, they were passed on to the next trial. There was no time limit for the response. Participants were tested individually in a quiet computer room.

3.0 Results – Pilot

The number of errors and response times were computed for each minimal pair and then averaged across all pairs of a given phonemic contrast for each participant. The results are summarized in Table 2.

Table 2 Average response times and proportion of errors made by adult English speakers.

		Average proportion of errors	*Average response time*
Singleton consonants	/p/ - /pʰ/	0.45	1538
	/b/ - /bʰ/	0.18	1425
	/k/ - /kʰ/	0.20	1435
	/g/ - /gʰ/	0.32	1560
	/t/ - /tʰ/	0.25	1433
	/d/ - /dʰ/	0.35	1388
	/ʈ/ - /ʈʰ/	0.20	1435
	/ɖ/ - /ɖʰ/	0.25	1500
	/tʃ/ - /tʃʰ/	**0.75**	**1629**
	/dʒ/ - /dʒʰ/	0.37	1446
	/ɽ/ - /ɽʰ/	0.05	1584
Geminates	/t/ - /t:/	0.2	1553
	/d/ - /d:/	0.05	1522
	/tʃ/ - /tʃ:/	0.1	1497
	/ʈ/ - /ʈ:/	0.05	1437
	/k/ - /k:/	0.1	1405

As can be seen in Table 2, the voiceless aspirated-unaspirated affricate pair /tʃ/-/tʃʰ/ **(in bold)** had the highest proportion of errors and the longest response times. This suggests that this pair was the most difficult for adult English speakers to discriminate.

4.0 Discussion - Pilot

The Urdu affricate contrast (/tʃ/-/tʃʰ/) presents a distinction that does not exist in English: Unaspirated /tʃ/ occurs in English but aspirated /tʃʰ/ does not. Urdu has four affricates, /tʃ/, tʃʰ/, /dʒ/ and /dʒʰ/, distinguished by aspiration and voicing, whereas the English affricates /tʃ/ vs. /dʒ/ are in principle distinguished by voicing only. To establish the acoustic similarities and differences between the English and Urdu affricates we conducted a comparative analysis of the Voice Onset Time (VOT) of the voiceless stops of English and the English and Urdu voiceless affricates (see Figure 1). Measurements were taken of the entire voiceless period, including, for the affricates, both the fricative energy for [ʃ] and the aspiration portion of /tʃʰ/. The onset of VOT was identified as the start of the release burst of /t/, and the offset as the start of periodicity in the following vowel. We analysed twenty tokens of word-initial /tʃ/ from each of two native speakers of British English and twenty tokens each of word-initial stops and affricates from each of the four native speakers of Urdu. The VOT shown in Figure 1 for syllable-initial English voiceless stops was taken from Docherty (1992: British English).

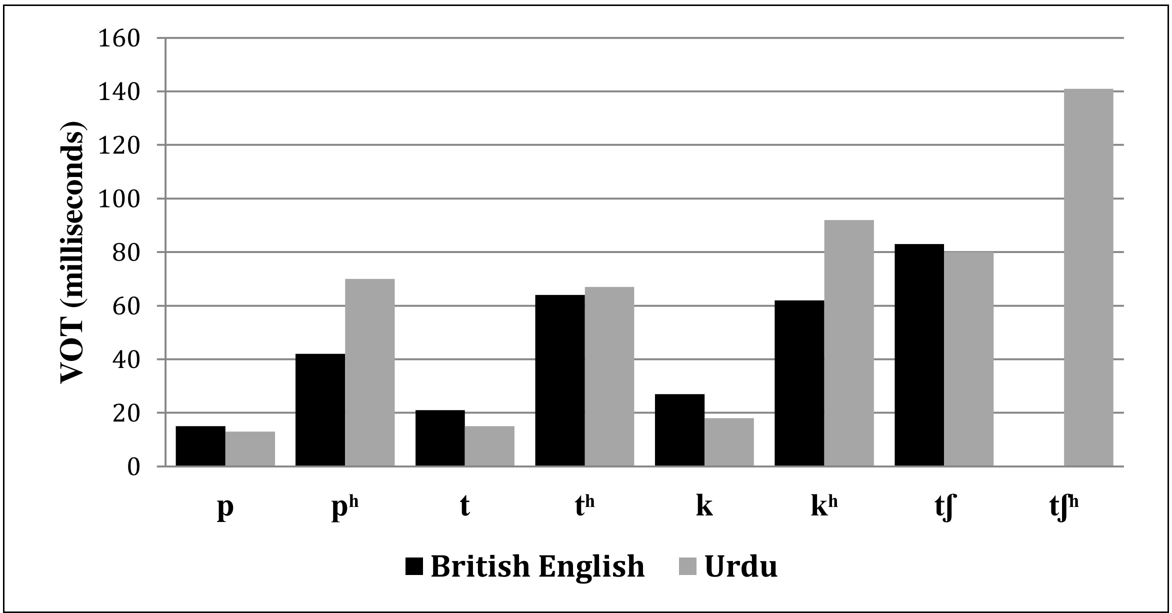


Figure 1 Voice Onset Times (ms) of English voiceless bilabial, dental, and velar stops (syllable-initial), Urdu voiceless bilabial, dental, and velar stops (word initial) and voiceless periods of Urdu and English voiceless affricates (word-initial).

The voiceless period for the English voiceless stops is around +40 to +80 ms and for the voiceless affricate it is around +80 (minimum: +63, maximum: +102, SD: 12.0); the voiceless period for the Urdu unaspirated affricate is around +80 (minimum: +32, maximum: +143, SD: 29.12) and for the aspirated affricate it is around +140 ms (minimum: +96, maximum: +185, SD: 24.34). Thus, the voiceless period of the English affricate is very similar to that of the Urdu unaspirated affricate but amounts to just over half that of the aspirated affricate of Urdu. This might account for the adult English speakers' lack of familiarity with and difficulty in discriminating the Urdu affricate contrast. Experiment 1 was designed to establish how English-learning infants respond to this unfamiliar aspirated-unaspirated contrast and to trace the expected change in their ability to discriminate these sounds as they approach their first birthday.

5.0 Experiment 1: Is there perceptual narrowing for a non-native aspiration contrast in affricates?

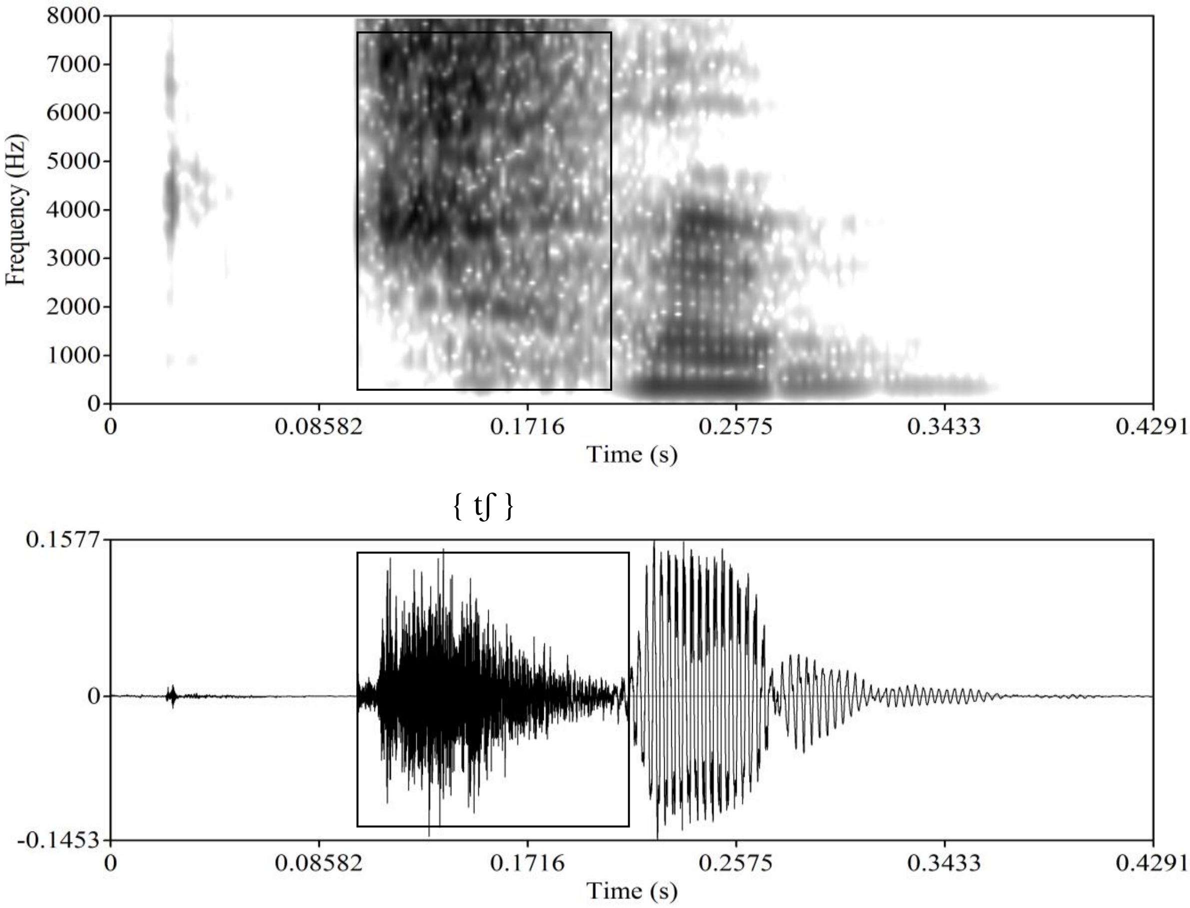
5.1 Experimental/Materials and methods – Experiment 1

5.1.1 Participants

Infants were recruited through advertisements in a local newspaper. Participants included 13 seven-month olds (mean age 210 days, range 204-217 days; 7 girls) and 16 eleven-month-olds (mean age 330.6 days, range 322-343 days; 7 girls). Only infants who were full term and without health problems were included in the experiment. An additional eight infants were excluded for fussiness and crying (7) or experimenter error (1). All infants were from monolingual English-speaking homes in York, England. None had any known hearing problem.

5.1.2 Stimuli

Twelve tokens each of the words /tʃʊp/ 'quiet' and /tʃʰʊp/ 'to hide' were recorded in a sound-attenuated recording room by a female native speaker of Urdu. The stimuli were presented to two other native speakers of Urdu for verification. An Urdu carrier sentence ‘can you say’ (تم یہ بولو) was used before each word. Spectograms and waveforms of a single example of each of the recorded Urdu words are shown in comparison with a corresponding English word, *chug* [tʃʊg] (in the Yorkshire accent) in Figures 2 and 3. There is considerable difference in the duration of aspirated and unaspirated affricates (Urdu voiceless aspirated affricate: 328 ms; voiceless unaspirated affricates, Urdu 254 ms, English, 265 ms). Also, the Urdu aspirated affricate in Figure 2 has relatively more intense frication (the affricate portion has been marked with brackets) than the unaspirated affricates of either Urdu (bottom part, Figure 2) or English (Figure 3).



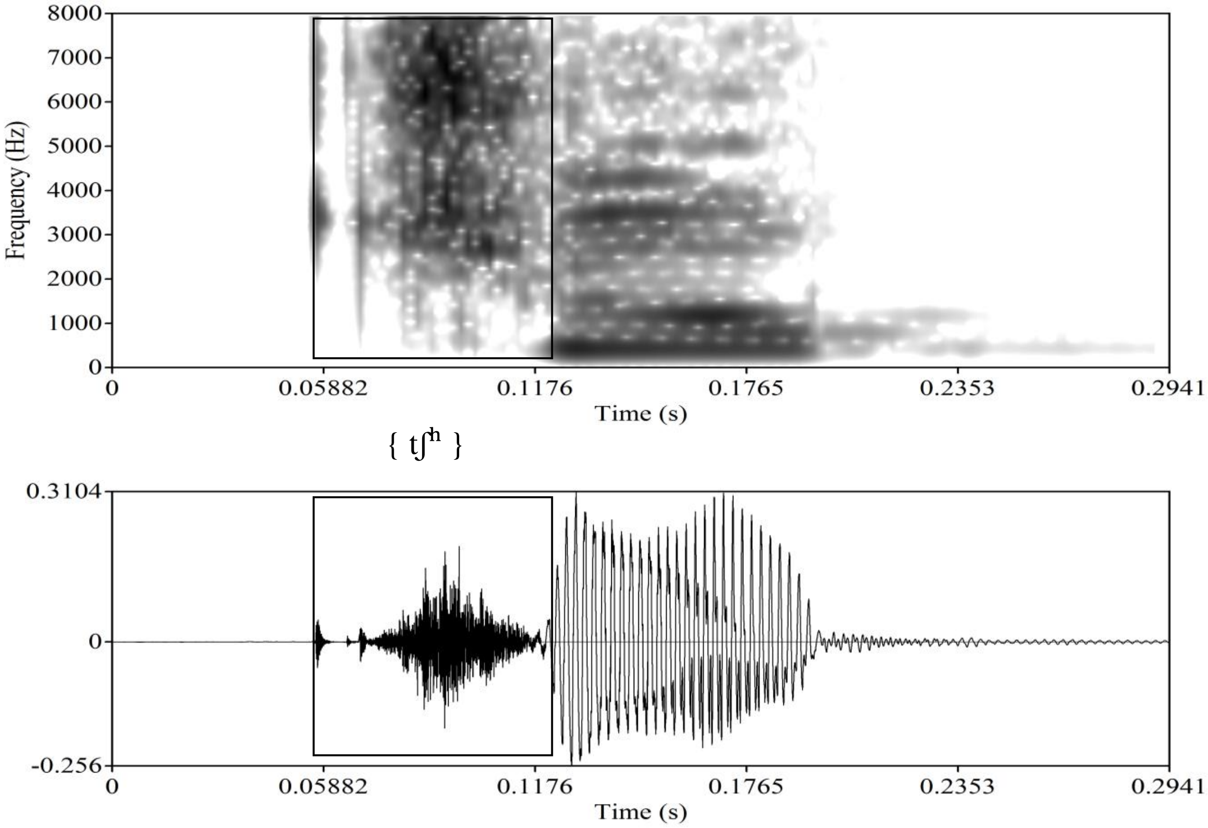


Figure 2 Spectrograms and waveforms of the Urdu words used in the study, /tʃʰup/ (top two panels) and /tʃup/ (bottom two panels). The two affricates are indicated by a box.

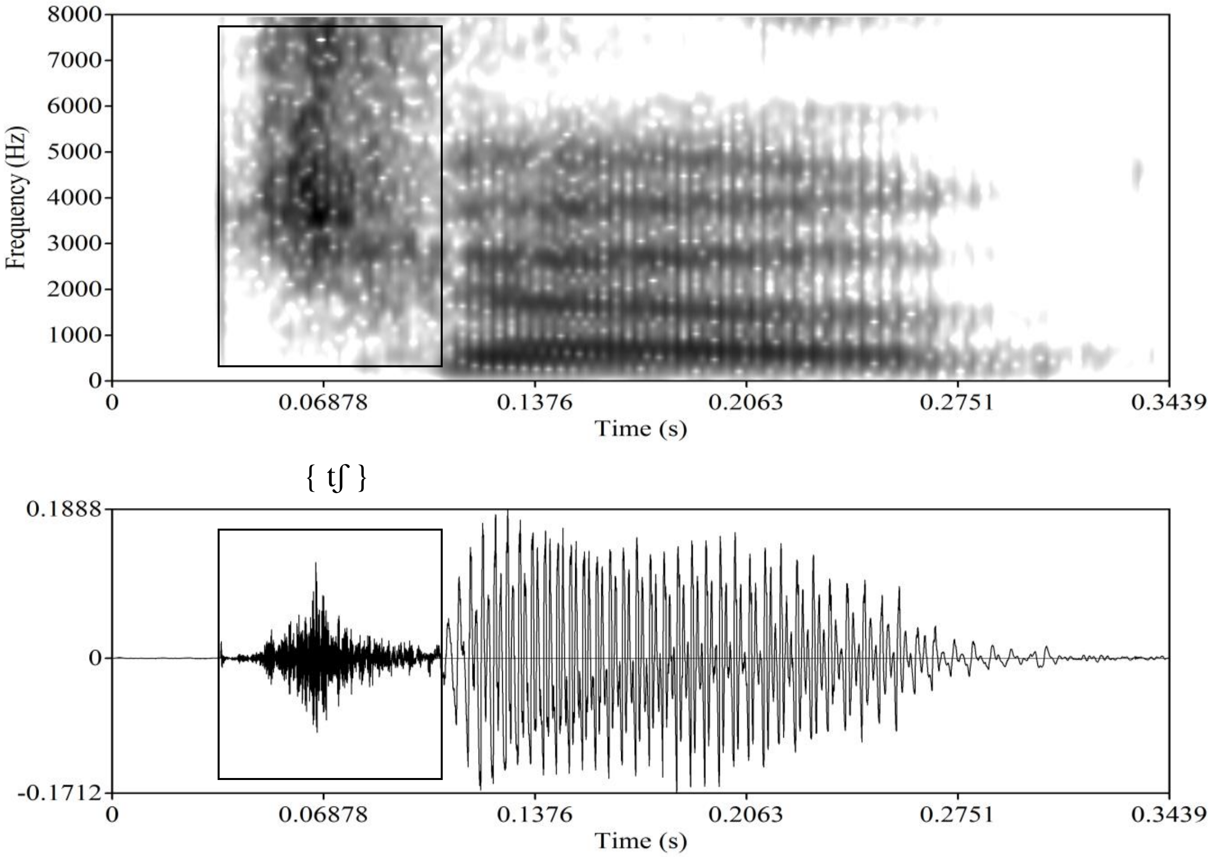


Figure 3 Spectrogram and waveform for English word *chug* /tʃʌg/. The affricate /tʃ/ is indicated by a box.

All tokens from this first recording were analysed acoustically for maximum amplitude, mean amplitude, mean F0, max F0, min F0, range F0 and duration, using Praat version 5.3.17. T-tests were carried out across all measures. The difference in the duration of tokens of the two words was statistically significant (p = .04) because the fricative portion of the aspirated affricates is necessarily longer than that of unaspirated affricates (Harris, Bell-Berti & Raphael, 1995, p. 161); no other significant differences were found. However, since near-significant differences were observed in F0 range between tokens (aspirated affricates had higher F0 values), we recorded another speaker. This second recording was also presented to two other native speakers of Urdu for verification. The analysis of this second recording again revealed near-significant differences for F0 range (p = .07), whereas no statistically significant differences were found for any other acoustic measures except duration. The near-significant difference in F0 range was therefore assumed to be an inherent property of the voiceless affricate pair in Urdu (see Table 3 for a full acoustic analysis). Six tokens of each of the two words that were the most similar acoustically (in maximum amplitude, mean amplitude, mean F0, pitch range, F0 and duration) were selected from the second recording to be used as stimuli.

Table 3 Acoustic measures of the voiceless aspirated/unaspirated affricate contrast /tʃʊp/ - /tʃʰʊp/ used in the experiment (from the second speaker). The table presents values averaged across the six tokens of each word.

		S1 - /tʃʊp/	*S2 - /tʃʰʊp/*
Max Amplitude (dB)	$\bar{x}$	*72.28*	*72.14*
SD		*1.56*	*0.71*
Mean Amplitude (dB)	$\bar{x}$	*69.61*	*69.30*
SD		*1.89*	*0.77*
Mean F0 (Hz)	$\bar{x}$	*293.71*	*293.57*
SD		*4.35*	*8.53*
Max F0 (Hz)	$\bar{x}$	*306.39*	*313.16*
SD		*4.05*	*8.37*
Min F0 (Hz)	$\bar{x}$	*281.02*	*281.21*
SD		*7.71*	*7.89*
Range F0 (Hz)	$\bar{x}$	*25.38*	*31.95*
SD		*3.89*	*7.07*
Duration (s)	$\bar{x}$	*0.250*	*0.468*
SD		*0.007*	*0.024*

5.1.3 Apparatus and Procedure

Testing took place in a dimly lit three-sided booth (120 x 122 cm) with black panels in a soundproof room. The stimuli were presented from a Yamaha KX-390 sound player through loudspeakers placed on both sides of the booth. The volume was adjusted with the help of a Tenma 72-6635 DP level meter. The infant was seated on the mother's lap approximately 45 inches from the monitor. The mother wore sound-cancelling Bose QC-15 headphones through which multi-talker babble created from the test stimuli was played to mask the auditory stimuli presented to the infants. Mothers also wore earplugs to enhance the masking.

An experimenter sat in the control room outside but adjacent to the soundproof room. Stimulus presentation was controlled by a Mac OSX 10.6.8. A Sony mini DV-HC27 video camera, hidden in the booth, recorded the infant and projected the footage onto a LCD Video Monitor XVIS8 in the control room, from which the experimenter could monitor the infant's looking behavior. To ensure that the acoustic stimuli were completely masked the experimenter wore headphones delivering the same masking sound as used for the parents.

The experiment had two phases; habituation and testing (the methodology for the experiment is diagrammed in Fig. 4). Only one stimulus (either /tʃʰʊp/ or /tʃʊp/) was presented in each phase. Six different tokens of each stimulus were placed on a loop and played repeatedly in both habituation and test phases. The sequence of the presentation of the stimuli was counterbalanced so that half of the infants were habituated to /tʃʊp/, the other half to /tʃʰʊp/. The inter-stimulus interval was 750 ms. The audio segments were presented at approximately 69 dB. Each trial began with a red light flashing on the monitor to attract the infant's attention. When the experimenter judged that the infant was looking at the screen, a key was pressed to deliver the visual stimulus, a black and white checkerboard, to the testing-room monitor. At the same time, the auditory stimuli began playing from the two loudspeakers on both sides of the booth. The loudspeakers were located at equal distances from the infant. For this reason, the sound seemed to surround the infant. Whenever the infant fixated the checkerboard, the experimenter pressed a button, releasing it only when the infant looked away. If the infant looked away for two seconds, the trial ended and a new trial began. Infant looking time was measured for the center look throughout the experiment.

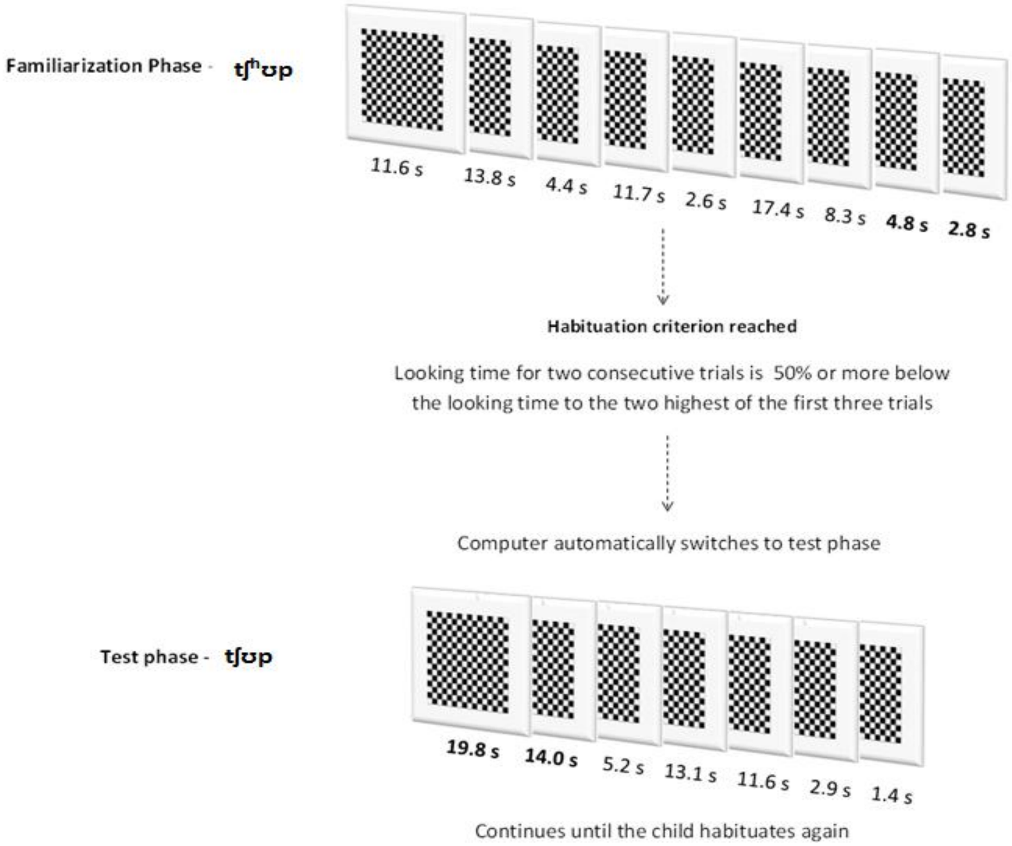


Figure 4 Summary of the habituation methodology used in the experiment. Each square represents a single trial, with time going from left to right. The looking times (in seconds) in the habituation and test phases are taken from one infant participant for illustrative purposes. The order of presentation of stimuli was reversed for half of the infants. Discrimination was measured by comparing the last two habituation trials and the first two test trials (in bold font). The infant whose times are shown here did discriminate, as there was a significant increase in looking times in the first two trials of the test phase.

Habituation was defined as two consecutive trials with fixation durations below 50% of the mean of the two highest of the first three trials (Pegg, Werker & McLeod, 1992). When the child reached the planned habituation criterion the computer automatically shifted to the contrasting stimulus for the test phase. Infants were expected to dishabituate in the test trial, showing an increase in looking time to the new stimulus, if they had discriminated the stimulus from the contrasting one presented in the habituation phase. For infants who failed to discriminate between the habituation and test stimuli no significant change in looking time was expected. Half of the infants listened to /tʃʊp/ in the habituation phase and /tʃʰʊp/ in the test phase and half heard the stimuli in the reverse order. The experimenter was unaware of the point at which the infant reached the habituation criterion. The number of habituation trials was not fixed in advance: Different infants received different numbers of habituation trials, depending upon the time they took to become habituated (with a range of 6-26 trials to habituate. The maximal possible number of trials to habituation was set at 40). The test phase continued until the infant habituated again (following Best et al., 1988; Best et al., 1995; Best & McRoberts, 2003). Maximal trial length was set at 30 seconds.

6.0 Analysis and results – Experiment 1

Discrimination was assessed by comparing mean looking time over the last two habituation trials (pre-shift phase) to mean looking time over the first two trials of the test phase (post-shift phase). A significant increase in mean looking time during the post-shift relative to the pre-shift phase is taken as evidence that the infant has detected the stimulus change. A discrimination value was calculated to minimize the effect of individual differences in looking times. This involved dividing the mean looking time in the first two test trials by the sum of the mean looking time in the first two test trials plus the mean looking time in the last two habituation trials. The point of no discrimination was set at 0.5 – in other words, equal looking in the two phases. A value over 0.5 indicates that the infant looked more towards the stimuli in the test phase, which signifies discrimination. A value below 0.5 indicates longer looking in the habituation phase, which means that the change was not detected in the test phase.

Discrimination values were calculated for each age group (see Fig. 5). An independent t-test showed no significant difference between the discrimination values of the two groups (t = 1.279, df = 27, p = 0.212, 2-tailed). Next, 1-tailed one-sample t-tests were run on each group to test for discrimination in that group against a maximal no-discrimination value of 0.5. The 7-month-olds showed increased looking in response to the test stimuli (M = 0.628, SD = 0.148, t = 3.117, df = 12; p = 0.005), whereas the 11-month-olds’discrimination score was not significantly above chance (M = 0.553, SD = 0.161, t = 1.326, df = 15, p = 0.103).

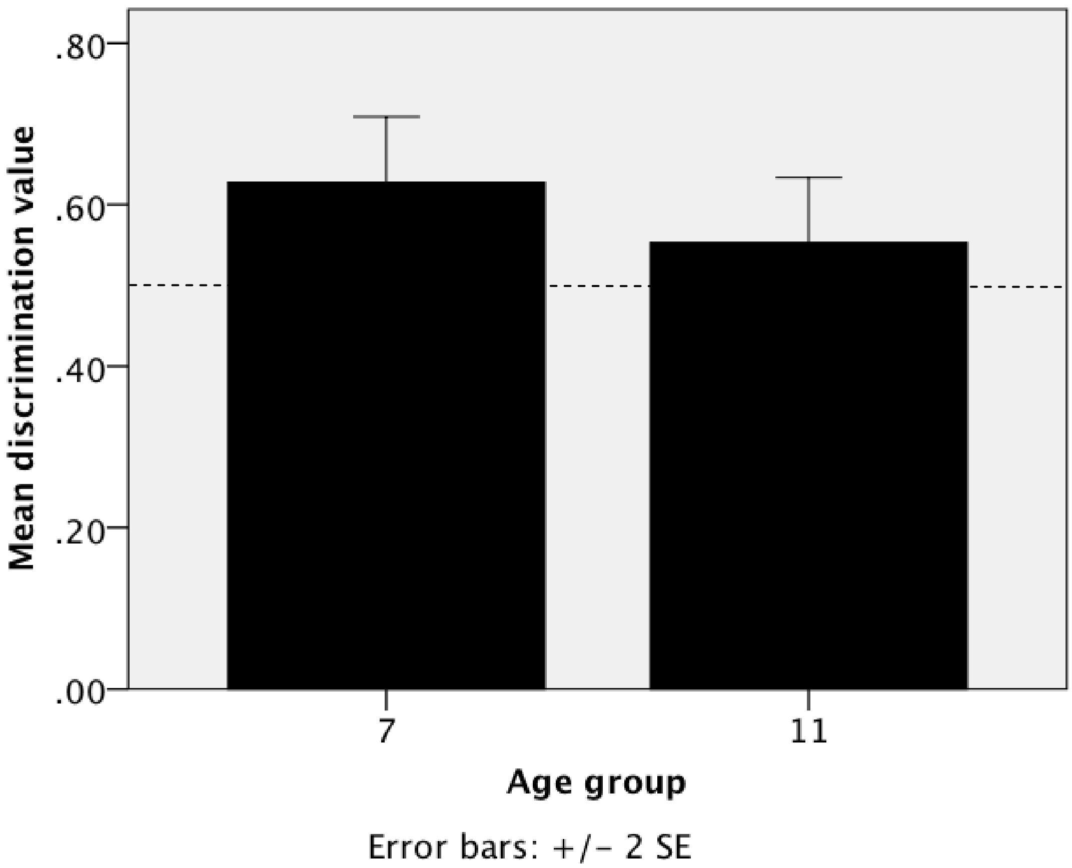


Figure 5 Discrimination values for 7- and 11- month olds. The reference line shows the point of no discrimination, 0.5. Error bars: +/- 1 SE.

Closer impressionistic inspection revealed that the order of presentation of stimuli affected the looking times of the older group of infants. Infants who heard the aspirated affricate first showed a longer looking time for the test stimuli, whereas infants who heard the unaspirated affricate first did not. To further investigate these possible order effects we ran exploratory one-sample 1-tailed t-tests on the discrimination scores of the subgroups of both younger and older infants against a maximal no-discrimination value of 0.5. (Note that we did not run an ANOVA on the results of Experiment 1 because here there was a single independent variable – age. We did not intend to look for order effects in this study, so did not use *order* as one of the independent variables. Our analyses regarding order effects in this experiment were all exploratory and post-hoc. Order effects were tested in a planned way in Experiments 2 and 3.)

Since the analyses relating to order effects were run after viewing the data had led us to notice what seemed like an asymmetry, the significance values calculated for these one-sample t-tests can only be taken to indicate possible avenues for future research; they do not signal significance in the usual sense. In the older group the mean discrimination value of the infants habituated to the unaspirated stimulus and then tested on the aspirated stimulus was 0.509 (SD = 0.168); this proved not to be significantly different from no discrimination (t = 0.157; df = 8, p = .440). Only four out of nine infants showed discrimination by having higher looking times in the test phase as compared to the habituation phase. The mean discrimination value for infants with the reverse order of presentation of stimuli was 0.611 (SD = 0.142); here a 'significant' effect was observed (t = 2.058, df = 6, p = 0.043): Five out of seven infants showed discrimination. Among the 7-month-olds, the mean discrimination value of those habituated to the unaspirated stimulus and tested on the aspirated stimulus was 0.631 (SD = 0.171; t = 2.030; df = 6, p = .045), with five out of seven infants showing discrimination; the mean discrimination value for those habituated to the aspirated stimulus and tested on the unaspirated stimulus was 0.623 (SD = 0.131; t = 2.307; df = 5; p = .035), with four out of six infants showing discrimination. The results hint that order of presentation of stimuli had no effect on the 7-month-olds (see Figure 6).

Two two-tailed independent t-tests (one for each age group) were run to examine whether there were differences in habituation times between the conditions. For the 7-month olds there was no significant difference: Mean habituation time to /tʃʰ/ was 111.34s, SD = 70.62 and to /tʃ/ it was 112.71s, SD = 48.89, t = -0.039, df = 11, p = .970. For 11-month-old infants the mean habituation time to the non-native-like /tʃʰ/, 66.60s, SD = 24.00 was significantly different from the habituation time to the more native-like /tʃ/, 102.27s, SD = 33.22, t = -2.389, df = 14, p = .032. The difference in the number of habituation trials for the two types of stimuli was not significant for either age group: For the 7-month-olds the average number of trials for habituation to /tʃʰ/ was 13.67, SD = 8.96, and to /tʃ/ it was 10.00, SD = 5.39, t = .911, df = 11, p = .382. For the 11-month-olds the mean number of trials for habituation to /tʃʰ/ was 8.71, SD = 2.75 and to /tʃ/ it was 9.11, SD = 1.96, t = -0.337, df = 14, p = .741.

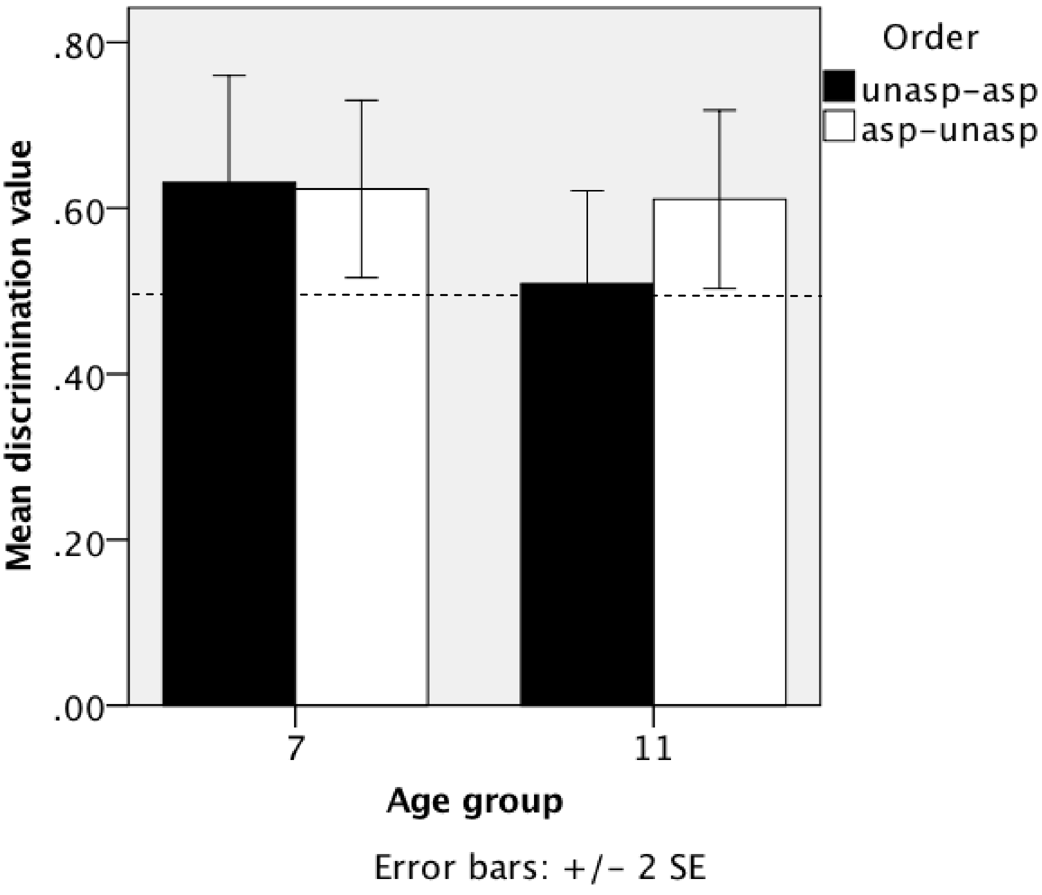


Figure 6 Discrimination values for both groups of infants by order of presentation of stimulus. . Error bars: +/- 1 SE.

7.0 Discussion

Experiment 1 was conducted to explore discrimination of a non-native aspiration contrast in affricates by infants from English-speaking homes. It was found that 7-month-olds successfully discriminated the contrast whereas the 11-month-olds did not. (The fact that the mean discrimination values of the two groups did not significantly differ is not evidence of a lack of perceptual narrowing; it merely shows that the difference is not large. The perceptual-narrowing claim is not about older infants being different from younger infants, but about whether younger infants succeed in discriminating a contrast that older infants do not discriminate; this was tested by comparing each group's discrimination value to 0.5.)

Order of presentation of the stimuli was not the focus of this experiment. Further exploration of the results showed that order of presentation did not affect the performance of the younger group of infants. However, a potential trend was observed in the older group: Infants showed better performance when the aspirated stimulus /tʃʰʊp/ was presented first. On the other hand, the subgroups tested were small (8 infants habituated to the unaspirated and tested on the aspirated stimuli and 7 infants in the subgroup given the opposite order), and asymmetry was not specifically targeted in this experiment and was only investigated after we had already seen the results

The habituation time differences between the two orders of presentations in the 11-month-old group may be spurious: Firstly, this finding is not mirrored in the analysis of number of trials. Secondly, habituation is *more rapid* to the less familiar segment. We looked at length of time to habituate as an additional measure of difficulty in processing the novel sound but, if anything, the novel sound seems to be processed more quickly by the 11-month olds. This finding is hard to explain at this stage.

The asymmetry hinted at in the results for our 11-month-old group suggests that the ability to distinguish non-native sounds may not have been lost in the older infants. Note that although a group size of 13-15 infants was suitable for testing perceptual decline in a group as a whole, it is insufficient for testing subgroups within each group. In addition, as mentioned above, the asymmetry we appeared to be seeing was not what we had set out to find. Accordingly, we decided to run Experiment 2, with larger subgroups at each age, in order to test specifically for order effects in English infants' discrimination of the non-native Urdu affricate contrast. We expected that, as before, the younger group would show discrimination regardless of order of presentation of the stimuli whereas the performance of the older group of infants would be affected by the order in which the stimuli are presented.

8.0 Experiment 2: Is there an order effect for the non-native aspiration contrast in Urdu voiceless affricates?

8.1 Experimental/Materials and methods – Experiment 2

8.1.1 Participants: Thirty 7-month olds (mean age 224 days, range 208-228; 17 girls) and thirty 11-month olds (Mean age 336 days, range 320-340 days; 14 girls) were recruited through advertisements in newspapers. Only full-term infants from monolingual homes were included in the experiment.

8.1.2 Stimuli: The same tokens of /tʃʊp/ and /tʃʰʊp/ were used as in Experiment 1.
8.1.3 Procedure: The procedure was identical to that used in Experiment 1.

9.0 Analysis and Results – Experiment 2

Both the 7- and the 11-month-old groups included subgroups of 15 infants each who received opposite orders of presentation. Figure 7 shows clearly that order of presentation of stimuli had no effect on the performance of the 7-month-olds. In contrast, an asymmetry was observed in the case of the older group; infants showed discrimination only when the stimuli with the aspirated affricate were presented first.

A preliminary two-tailed t-test examined the difference in habituation times between conditions. No significant difference was found in either age group: 7-month-olds' mean habituation time to /tʃʰ/ was 101.94s, SD = 55.96, and to /tʃ/ it was 87.90s, SD = 40.18, t = 0.789, df = 28, p = .437. For 11-month-olds mean habituation time to /tʃʰ/ was 91.15s, SD = 46.54 and to /tʃ/ it was 79.22s, SD = 46.63, t = 0.702, df = 28 p = .489. The difference in the number of habituation trials was not significant for either age group either: For 7-month-olds the mean number of trials for habituation to /tʃʰ/was 10.20, SD = 4.65, to /tʃ/ it was 9.00, SD = 3.76, t = 0.777, df = 28, p = .443. For 11-month-olds the mean number of trials for habituation to /tʃʰ/ was 9.40, SD = 4.29, and to /tʃ/ it was 9.33, SD = 5.95, t = -0.282, p = .780). (Note that in Experiment 2, as with the 7-month-old group in Experiment 1, we found no difference in habituation time or number of trials to habituation between infants exposed to the different orders. There is thus no indication that one of the affricates is inherently, or acoustically, more attention-grabbing and therefore slower to lead to habituation.)

An independent two-way ANOVA with age (2 levels: 7, 11 months) and order (2 levels: aspirated-unaspirated, unaspirated-aspirated) as the independent variables was run with discrimination values as the dependent variable. The main effect of age was not significant (df = 1; $F = 2.984$; p = .09). The main effect of order was significant (df = 1; $F = 16.360$; $p < 0.001$), with aspirated-unaspirated resulting in significantly higher discrimination values (M = .704) than unaspirated-aspirated (M = .591). The interaction between age and order was also significant (df = 1; $F = 10.494$; $p < 0.01$). To further investigate the interaction, we followed this ANOVA with a pair of independent t-tests (using the Bonferroni correction), one on each age group. For the 11-month olds there was a significant difference between the two orders ($t = 4.326$, df = 28, $p < 0.001$, 2-tailed), with aspirated-unaspirated (M = .725; SD = 0.107) showing a higher discrimination value than unaspirated-aspirated (M = .522; SD = 0.147). However, no significant difference was found between the two orders with the 7-month-olds (for aspirated-unaspirated order: M = 0.683, SD = 0.079; for unaspirated-aspirated: M = 0.661, SD = 0.086, $t = 0.746$, df = 28, p = .462, 2-tailed). Figure 7 shows clearly that the order of presentation of the stimuli affected only the 11-month-olds; infants showed discrimination only when the stimuli with the aspirated affricate were presented first.

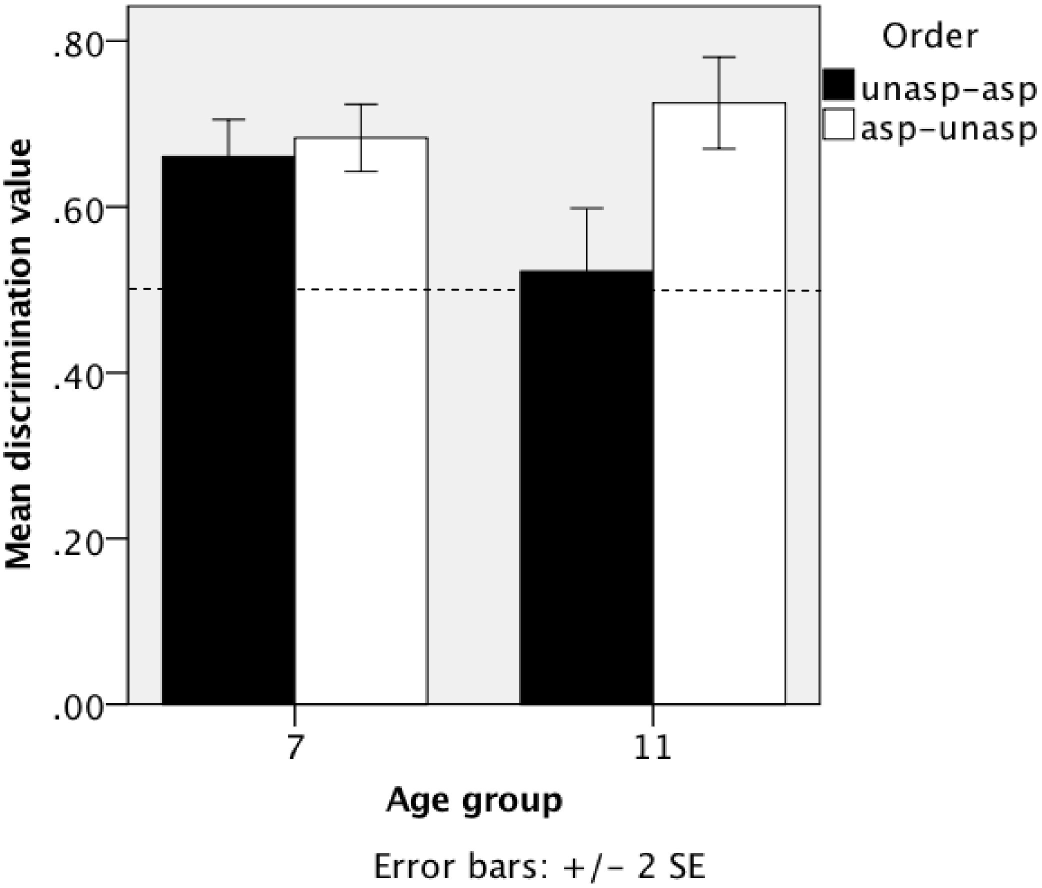


Figure 7 Discrimination values of 7- and 11-month olds by order of presentation of stimulus. Error bars: +/- 1 SE.

Note that the absence of a main effect of age is due to the fact that both age groups (when the two orders of presentation are grouped together) showed similar discrimination for the test stimuli. Two one-tailed one-sample t-tests (with Bonferroni corrections) were run to see whether these discrimination values are significantly higher than no discrimination (0.5). The 7-month-olds showed a significant increase in looking times in response to the test stimuli (M = 0.672, SD = 0.082, t = 11.501, df = 29, $p <$ 0.001) as did the 11-month-olds (M = 0.624, SD = 0.163, t = 4.146, df = 29, $p < 0.001$). Finally, we ran four 1-tailed one-sample t-tests, to test whether each of the subgroups' discrimination value was significantly higher than no discrimination (0.5). Both subgroups of 7-months olds showed discrimination: for the aspirated-unaspirated order: t=9.001, df = 14, p < .001 (15 out of 15 infants showed discrimination); for unaspirated-aspirated: t = 7.233, df = 14, p < .001 (14 out of 15 showed discrimination). Among the 11-month-olds the subgroup tested with the aspirated-unaspirated order showed discrimination: t = 8.164, df = 14, p < .001 (14 out of 15 infants showed discrimination), but the subgroup tested with the unaspirated-aspirated order did not: t = 0.578, df = 15, p = .286 (only nine out of 15 infants showed discrimination). In total about two-thirds of the infants in the older group (23 out of 30) showed discrimination, which explains the overall significant results for this group. Thus in Experiment 2 no evidence of perceptual narrowing was observed for the older group taken as a whole, but one of the subgroups, namely, the one tested with the unaspirated-aspirated order, does show perceptual narrowing.

10.0 Discussion

Experiment 2 was conducted to explore the discrimination of a non-native aspiration contrast in affricates by infants from English-speaking homes. The order of presentation of the stimuli did not affect the performance of the younger group of infants. However, the 11-month-olds showed better performance when the aspirated stimulus /tʃʰʊp/ was presented first. This suggests that the ability to distinguish non-native sounds had not been lost in the older infants. Note, however, that when testing a single group of 11-month-olds in Experiment 1, we did find loss of the ability to discriminate the contrast. Only when we ran a group twice as big as that used in Experiment 1 and in other similar experiments (e.g., Werker & Tees, 2002; Best & McRoberts, 2003) did older infants show the ability to discriminate at the group level.

In Experiment 2, the subgroup habituated to the aspirated affricate behaved similarly to the 7-month-olds, showing no signs of developmental decline, unlike the other subgroup, which was habituated to the native-like aspirated affricate. The results make the evidence for perceptual narrowing more complex, since the older group of infants showed signs of developmental decline when presented with the stimuli in one order but not in the other, with the very same contrast. This suggests that the perceptual narrowing observed in infants at the end of first year may depend upon additional factors beyond those considered so far. At the end of the first year infants tend to show a decline in the perception of non-native contrasts that are not functional in the ambient language. However, if at that age infants *can* show discrimination for the non-native contrast, without special training, when presented with the stimuli in a specific order, what does that mean for the finding of perceptual narrowing? Is perceptual narrowing merely a function of task characteristics? Or do the order effects found here tell us something about infants' sound representations? We will come back to this in the main discussion. But first we test whether adult English speakers show insensitivity to the contrast /tʃʰ/ - /tʃ/ and whether this 'narrowing' of their perception is advanced to such an extent that they cannot discriminate between the two sounds, even when presented with the aspirated affricate first. In order to investigate that issue we compared the performance of adult English speakers to a group of native Urdu-speaking adults.

11.0 Experiment 3: Is the asymmetry in discrimination for non-native consonants maintained in adulthood?

11.1 Experimental/Materials and methods – Experiment 3

11.1.1 Participants: Twenty English-speaking adults (18 British, 2 Americans) were recruited through word of mouth, advertising to students and staff at the Department of Language and Linguistic Science at the University of York and through social media. All participants were born and brought up in monolingual English-speaking homes and were studying at the University of York at the time of the experiment. The mean age was 26 (range 18 – 35 years). Twenty adult native Urdu speakers were recruited in Pakistan through word of mouth, acquaintances and social media. All were born and raised in Pakistan and ranged in occupation from postgraduate students to working professionals. The mean age was 30 (range 25-37 years).

11.1.2 Stimuli: The stimuli consisted of 12 minimal pairs of Urdu words, one of which contains the phoneme /tʃ/ and the other the phoneme /tʃʰ/ as word onset.[4] The stimuli were recorded by a female adult native-Urdu speaker. Each of the words was recorded within the Urdu carrier sentence 'can you say X'. Each word was recorded three times, resulting in three different tokens of each word. We created 24 'different' word pairs (minimal pairs) – 12 with the word containing the aspirated segment first and 12 with the word containing the unaspirated segment first. We also created 24 'same' pairs, 12 with two different tokens of the same word, both including the unaspirated segment, and 12 with both tokens including the aspirated segment (see Appendix for word lists). No token was used more than once.

11.1.3 Procedure: The procedure was identical to that used in the adult pilot test. The participants were tested using an AX discrimination task identical to that used in the adult pilot.

12.0 Analysis and results

To compare the English (non-native) adult listeners to the Urdu (native) adult listeners we ran two mixed ANOVAs, with native language (2 levels: Urdu, English) as a between-participant variable, and same-different (2 levels: same, different) and order (2 levels: aspirated first, unaspirated first) as within-participant variables. The dependent variables were proportion of correct responses (out of 12) in the first ANOVA and reaction time (RT, for correct responses only) in the second. The responses of one of the Urdu participants were taken out of the analysis, because the pattern of their responses showed that they had not engaged with the task (they responded 'same' to 47 out of 48 trials). The final sample therefore included 19 Urdu participants and 20 English participants. In the ANOVA run on proportion of correct responses there was no main effect of order ($F(1) = 0.038$, p = .847) nor of same-different ($F(1) = 0.060$, p = .807). There was, however, a significant main effect of native language, with the mean proportion of correct responses for the Urdu native listeners being higher (M = .871, SD = .106) than that of the English listeners (M = .694, SD = .070): df = 1; $F(1) = 38.022$; $p < .001$.). No interaction was significant.

We followed the ANOVA with four two-tailed one-sample t-tests on the proportion of correct responses for each trial type, separately for each native-language group (using the Bonferroni correction), to assess whether each group's performance on each type of stimulus was significantly different from chance (0.5). Both groups performed significantly better than chance on all types of trials: For the Urdu adults: Different-Aspirated first: M = .833, SD = .171, $t = 8.485$, df = 18, $p < .001$, Same-Aspirated first: M = .900, SD = .149, t = 11.716, df = 18, p < .001, Different-Unaspirated first: M = .873, SD = .156, t = 10.445, df = 18, p < .001; Same-Unaspirated first: M = .877, SD = .109, t = 15.098, df = 18, p < .001. For the English adults: Different-Aspirated first: M = .708, SD = .147, $t = 6.345$, df = 19, $p < .001$, Same-Aspirated first: M = .679, SD = .158, t = 5.062, df = 19, p < .001, Different-Unaspirated first: M = .725, SD = .151, t = 6.674, df = 19, p < .001; Same-Unaspirated first: M = .663, SD = .136, t = 5.325, df = 19, p < .001. We ran a d' analysis to assess the sensitivity of the discrimination in each of the groups. This type of analysis combines the proportion of *hits* (i.e., 'different' trials receiving correct 'different' judgments from the participants) with the proportion

[4] Four additional pairs were used in the test but were later remove from the analyses, due to one member of the pair being a word and the other a nonword.

of *false alarms* (i.e., 'same' trials receiving erroneous 'different' judgments from the participants) (Keating, 2005). The d' analysis shows higher sensitivity on the part of the Urdu adults (a mean of 1.73 in the Aspirated-first condition and 1.75 in the Unaspirated-first condition) than the English adults (a mean of 1.57, Aspirated-first, and 1.60, Unaspirated-first). Order of presentation has little if any effect in either group.

In the ANOVA run on RTs for correct responses there was a main effect of native language, with Urdu native participants (M = 1188.89 ms, SD = 308.96) responding faster than English participants (M = 1559.58, SD = 437.16): $F(1) = 9.260$, $p = .004$) and of same/different, with 'same' trials receiving slower responses (M = 1436.90, SD = 486.26) than 'different' trials (M = 1321.08, SD = 419.63): $F(1) = 4.221$; $p = .047$). None of the interactions were significant.

12.1 Discussion

Experiment 3 showed that native English adults do indeed show reduced sensitivity for the Urdu aspirated-unaspirated affricate contrast: they were both less accurate at discriminating the contrast and slower to respond to the stimuli than Urdu-speaking adults. Although they found the task difficult (as evidenced by their slow responses), English adults were still able to discriminate the two affricates at an above-chance level. This result, however, is not unlike findings regarding poor (but above chance) discrimination for other non-native sounds in adult listeners: Japanese listeners tested on the perception of English /la/-/ra/ identified the correct phoneme around 70% of the time (/r/ was identified correctly on 71% of trials and /l/ on 67% of trials: Hattori & Iverson, 2009). Low proficiency English participants, native speakers of Saudi Arabic, tested on perception of English phonemes identified /p/ correctly 74% of the time and /b/ 68% of the time (Alshangiti, 2015). When we set out to test the English adults we expected to find their performance to be poorer than it actually was, and we expected that we might see an asymmetry. As it happens, adults performed well enough on both orders; there was no evidence for asymmetry. However, it is possible that the AX task that we used was not difficult (and therefore sensitive) enough, and that a harder task, which would have resulted in reduced performance among the English adults, would have shown the advantage of one order of presentation over the other.

As regards the difference in performance between 'same' and 'different' pairs in the RT analysis, we take this to be task-dependent: Since every pair used for judging included two different tokens, the members of the pair were always to some extent different from one another. It is therefore arguably easier for a listener to respond to this difference with a 'different' judgment than to overlook small differences between tokens and judge them the 'same'. The d' analysis actually takes both of these types of trial into account and combines information about accuracy to give a sensitivity score. However, the issue of the different task demands in the trials involving similar vs. different stimuli is not directly relevant to the developmental question we are investigating here.

The results from the adult study should be interpreted with caution. First, the infants and adults were tested on different tasks and with different stimuli. It is also possible that the relatively short Inter Stimulus Interval (ISI) that we used (300 ms) favours participants using auditory rather than phonetic processing. Repp and Crowder (1990) pointed out that with a short ISI listeners rely on continuous auditory information for comparison or identification, while with an increase in ISI, listeners rely on phonetic information or category labels. Our results might have looked different, therefore, had we used a longer ISI. However, Tsushima et al. (2003) find asymmetry in consonant discrimination by adults even at very short ISI (100 ms), though only marginally at an ISI of 300 ms: These researchers tested Japanese adults on /b/-/v/ in an AX discrimination task. They found order effects, such that discrimination is better when /v/ is presented first than when /b/ was first (only /b/ occurs as a phoneme in Japanese). Therefore, whether the ISI used here was a problem or not remains unclear. Secondly, a considerable difference was found in adult performance between the pilot and Experiment 3. This difference can be interpreted in the light of other adult discrimination studies showing the effects of task familiarity on discrimination (Tsushima et al., 2003, 2005; Tsushima 2007, 2011). In Tsushima et al. (2003), order effects in the expected direction were systematically observed only in the pretest, which is comparable to our Experiment 3, but not after repeated training with the same stimuli over several days. The authors attributed the disappearance of the order effects after the pretest to participants' increased proficiency at discriminating the contrast. In Tsushima (2011) Japanese adults were again tested on the /b/-/v/ stimuli, using a fixed category procedure (for half of the listeners /b/ always occurred first and vice versa). It was found that the participants in the /b/-first group were able to take advantage of the frequent presentation of /b/ as the first stimulus by picking up critical acoustic cues that helped in discrimination - and that are also used in the native language. Due to their unfamiliarity with the acoustic properties of /v/ the adults in the /v/-first group could not similarly gain from the repeated presentations. The Tsushima (2011) study found that with increased familiarity, this order effect not only disappeared, but was reversed (see Tsushima, 2007, for similar results for /l-/r/). In our study the English adults in Experiment 3 listened to 48 pairs of minimal pairs featuring /tʃ/-/tʃʰ/, those in the pilot to only 2 pairs. It is possible that in Experiment 3 the discrimination of English-speaking adults improved as their increased familiarity with the stimuli increased, leading to relatively high performance and a loss of the order effect.

13.0 General Discussion

The main goal of the study was to investigate asymmetry in English infants' (7- and 11-month-olds) and adults' discrimination of a non-native Urdu contrast. Experiment 2 showed that the order of presentation had no effect on 7-month-olds. However, 11-month-olds discriminated successfully only when the aspirated affricate was presented first. These results confirm the existence of an order effect for the older group of infants. As no order effect was found for the younger group, we can conclude that the asymmetry found in this study is not due to a universal perception bias but must be a consequence of learning from the input. The asymmetry for consonant perception (observed in Experiment 2) was not expected, since no previous studies that we are aware of have reported such a learning-based asymmetry for consonants (apart from Segal et al., which found a similar but non-significant trend).[5] Based on previously published findings, we expected (in Experiment 1) to see a developmental decline for in discrimination of two non-native consonants that differ in Category Goodness, one being a good exemplar of a native sound category and the other a deviant one (based on

[5] Note that Mugitani et al. (2009) found a learning-based asymmetry for vowels. Interestingly, in that study, 18-month-olds showed an asymmetry in discriminating vowel length only when length was phonemically contrastive in their native language (Japanese). English-learning 18-month olds showed discrimination, regardless of order of presentation, even though vowel length is not contrastive in English. As such, this could be a case of perceptual narrowing in the opposite direction to that usually found: distinctions that are contrastive in the native language are not recognised, while those which are not contrastive are recognised.

Best's 1993 taxonomy; see Kuhl , 2004; Kuhl et al., 2008; 2008; Werker et al., 1981, Werker & Tees, 1983, 1984). Our findings from both Experiments 1 and 2 suggest that the finding of a decline in perception may depend, at least in part, on the order in which stimuli are presented, so that the decline may be a more nuanced phenomenon than has generally been assumed.

As discussed in the introduction, infants' early experience with language input plays a vital role in shaping their perceptual development. If two sounds do not contrast in infants' language environment, the lack of perceptual experience with the contrast attenuates infants' ability to recognize it. This attenuation makes the perception of stimuli in the region of that particular perceptual boundary less discriminable. At seven months of age, prior to the time when infants' perceptual development becomes attuned to the phonological categories of the ambient language, the infants in this study were able to discriminate the Urdu contrast. As the aspirated-unaspirated affricate contrast does not occur in English and infants are exposed only to unaspirated affricates, the input to English-learning infants likely has a unimodal distribution in this area, which leads to the formation, towards the end of the first year, of a broad single category for voiceless affricates rather than two separate categories. This native-category learning can make the discrimination of the non-native aspirated-unaspirated affricate contrast more difficult: The two phones fall within a single category for English (voiceless alveolar affricate) and differ in Category Goodness. Experiment 1 indeed showed the decline in discrimination of the contrast in 11-month-olds as a group. (Although 11-month-old infants in Experiment 2 did, as a group, discriminate the contrast, group sizes were much larger than is standard for this type of experiment, where more typically N = 15 or fewer in each age group, as in, e.g., Werker & Tees, 2002 and Best & McRoberts, 2003). Experiment 3 showed that native English adults exhibit a perceptual decline in, rather than a complete loss of, the ability to discriminate the two sounds in adulthood.

However, Experiment 2 showed that this picture may be too simple: At 11 months evidence for perceptual narrowing may or may not be found, depending not only on the child and his/her developmental stage but also on task characteristics. The 11-month-old infants showed better discrimination when the Urdu voiceless aspirated affricate – sufficiently different from its closest English equivalent to stand out for them – was played first. This finding is consistent with the perceptual magnet effect previously observed in within-category vowel discrimination by infants and adults, where more prototypical vowels act as perceptual magnets (Kuhl, 2004). Frequency of occurrence in the input plays an important role in shaping infant's perceptual categories (recall the bimodal/unimodal effect in Maye et al., 2002 and 2008). The voiceless period for English /tʃ/ (+83 ms) is very close to that of the Urdu voiceless unaspirated affricate (+80 ms) but is half the duration of the Urdu voiceless aspirated affricate /tʃʰ/ (+140 ms). In terms of frequency of exposure, then, English-learning infants would have heard many affricates similar to the unaspirated Urdu affricates but few if any affricates similar to the Urdu aspirated one. When the 11-month-olds heard the familiar or prototypical (unaspirated) affricate in the habituation phase, this may have activated various familiar exemplars, since English has many words starting with /tʃ/; this should result in strong activation of that phonetic category. Arguably as a result of this, when the non-prototypical affricate /tʃʰ/ was played after the prototypical affricate /tʃ/ it was assimilated to that category, blocking discrimination. On the other hand, when the unfamiliar or non-prototypical (aspirated) affricate was played first, it likely failed to activate any familiar exemplars very strongly or it may have activated exemplars of different kinds, belonging to no one category. The infants would have been unable to relate it straightforwardly to anything they had heard before. Thus it presented a sharp contrast to the native-like affricate /tʃ/ that followed, facilitating discrimination of the test stimuli.

Although *categorical perception* is assumed to obtain for consonants – as opposed to gradient perception for vowels – there is evidence against strong categorical perception in consonants (as discussed in the Introduction). Moreover, fluent speech contains a mix of central exemplars (prototypes) and not-so-typical exemplars, such that consonants may vary with speaker, context and co-articulatory effects (Howell, 1983; Nartey, 1984; Nittrouer & Studdert-Kennedy, 1987; Guenther, 1995; Suzuki, Kitamura, Masaki & Michi, 2001; Kleber, Harrington, & Reubold, 2011). In particular, VOT varies for stops in word-initial position as a function of the following vowel and listeners have to adjust their VOT boundaries in relation to this interaction (MacKain, 1982). Inference and prediction must play an important role in word or phoneme identification, given that listeners typically have no problem recognizing phonemes, even if the speakers fail to generate all the expected attributes (Medin & Barsalou, 1987). A clearly produced /b/ and a poorly produced /b/ will both be perceived as /b/ due to this interplay of inference and prediction. Exemplars of many consonantal categories vary along a critical continuum (like VOT), with the category membership of exemplars near the boundary being less strong and clear than that of exemplars close to the center or prototypes (see Grieser & Kuhl [1989] for a detailed discussion of how infants organize speech categories around [vowel] prototypes). Infants might react differently to consonant tokens near to vs. far away from the center of a category. Experiment 2 provides evidence in this regard by showing that the consonants presenting prototypical values acted as perceptual magnets in a perception test, assimilating neighboring exemplars. In contrast, consonants at the extreme end of the range, when presented first, aided discrimination.

Based on the three studies reported here the results appear to reflect a learning-based, language-specific effect, not a universal bias. The 7-month-olds showed no order effect for the non-native affricate contrast; only the 11-month-olds showed such an effect. No difference was found in the habituation time or number of trials to habituation between infants exposed to the two different orders in either age group in Experiment 2. This might suggest that there is nothing inherent to the affricate contrast to make one member acoustically more salient and thus easier for infants to habituate to. Moreover, in Experiment 3 with adults no order effects were observed in either the native Urdu or the English group, contrary to what would be expected in relation to a universal bias, although we realize that the adult study may have been insufficiently sensitive to capture any asymmetry. In future work it would be important to test a native-language control group, i.e., Urdu infants, on the Urdu affricate pair.

Asymmetries have been reported in a number of non-linguistic stimulus domains, such as line orientation and numbers (Rosch, 1975) and geometric figures and country concepts (Tversky & Gati, 1978), but no study other than that of Kuhl et al. (2006) has reported significant evidence for asymmetry in non-native consonant perception in infants. Kuhl et al. (2006) found asymmetries in consonant perception for /l/ and /r/ when testing Japanese and English infants. However, that study found asymmetry for both younger and older infants, regardless of native language. Their findings cannot, therefore, be traced to native-language learning leading to the attenuation of sensitivity to non-native contrasts, and accordingly they are not directly relevant to the issue of perceptual narrowing. They are more in line with the Polka & Bohn (2011) model of perceptual asymmetry due to universal perceptual biases. Our results and those of Segal et al. (2016) are the only ones we are aware of that show asymmetry in consonant perception under conditions in which asymmetry is seen in the older but not in the younger group tested. Table 2 compares the findings of Kuhl et al. (2006), Segal et al. (2016) and the present study.

Table 4 Comparison of present study findings with other studies on non-native consonant asymmetry

	Manner of articulation	*Direction of asymmetry*	*Asymmetry observed for*			
			Native[a]		*Non-native*	
			Younger	*Older*	*Younger*	*Older*
Kuhl et al. (2006)	sonorants	/la/ - /ra/	Yes	Yes	Yes	Yes
Segal et al. (2016)	obstruents	/pa/ - /ba/	No	No	No	Yes
Dar et al. (present study)	obstruents	/tʃʰ/ /tʃ/	--	--	No	Yes

[a] -- indicates variable was not tested

14.0 Conclusions

The findings of the present study do not challenge studies showing perceptual narrowing in infants. Rather, the results suggest that each consonantal contrast may have its own developmental story, and that the narrowing observed in infant speech perception tasks depends to some extent on the particular task and the particular contrast.

A number of interesting questions remain. In the present study the presentation of /tʃʰ/ in the first phase aided the English-learning infants' discrimination, but can similar results be obtained with infants from Urdu-speaking homes? It remains for future studies to test Urdu-learning infants on the affricate contrast /tʃ/ - /tʃʰ/. The age effect is yet another issue: Are the asymmetries maintained at later stages for the contrasts that do not become functional in the native language? If not, at what age do they disappear? The findings of an order-effect in infant non-native consonant discrimination opens up new lines of research, which may shed new light on adult as well as infant processing of consonants.

Reference List

Alshangiti, W. (2015). *Speech production and perception in adult Arabic learners of English: A comparative study of the role of production and perception training in the acquisition of British English vowels* (Doctoral dissertation). Retrieved from http://discovery.ucl.ac.uk/1466643/1/Thesis.pdf.

Aslin, R. , & Pisoni, D. (1980). Some developmental processes in speech perception. In G. Yeni-Komshian, J. Kavanagh, & C. Ferguson, (Eds.), *Child Phonology, 2: Perception* (pp. 67-95). New York: Academic Press.

Altvater-Mackensen, N., & Fikkert, P. (2010). The acquisition of the stop-fricative contrast in perception and production. *Lingua*, *120*(8), 1898-1909. doi:10.1016/j.lingua.2010.02.010.

Best, C. T. (1993). Emergence of Language-Specific Constraints in Perception of Non-Native Speech: A Window on Early Phonological Development. In B. de Boysson-Bardies, S. de Schonen, P. Jusczyk, P. McNeilage, & J. Morton (Eds.), *Developmental Neurocognition: Speech and Face Processing in the First Year of Life* (pp. 289–304). Dordrecht: Kluwer Academic Publishers. doi.org/10.1007/978-94-015-8234-6_24.

Best, C., & McRoberts, G. (2003). Infant Perception of Non-Native Consonant Contrasts that Adults Assimilate in Different Ways. *Language and Speech*, *46*(2-3),183-216. doi:10.1177/00238309030460020701.

Best, C., McRoberts, G., LaFleur, R., & Silver-Isenstadt, J. (1995). Divergent developmental patterns for infants' perception of two nonnative consonant contrasts. *Infant Behavior and Development*, *18*(3), 339-350. doi:10.1016/0163-6383(95)90022-5.

Best, C., McRoberts, G., & Sithole, N. (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*(3), 345-360. http://dx.doi.org/10.1037//0096-1523.14.3.345.

Bohn, O., & Polka, L. (2001). Target spectral, dynamic spectral, and duration cues in infant perception of German vowels. *Journal of the Acoustical Society of America*, *110*(1), 504-515. doi:10.1121/1.1380415.

Bosch, L., & Sebastian-Galles, N. (2003). Simultaneous Bilingualism and the Perception of a Language-Specific Vowel Contrast in the First Year of Life. *Language and Speech*, *46*(2-3), 217-243. http://dx.doi.org/10.1177/00238309030460020801.

Carney, A., Widin, G., & Viemeister, N. (1977). Noncategorical perception of stop consonants differing in VOT. *Journal of the Acoustical Society of America*, *62*(4), 961-970. doi:10.1121/1.381590.

Conboy, B., Sommerville, J., & Kuhl, P. (2008). Cognitive control factors in speech perception at 11 months. *Developmental Psychology*, *44*(5), 1505-1512. doi:10.1037/a0012975.

Cowan, N., & Morse, P. (1986). The use of auditory and phonetic memory in vowel discrimination. *Journal of the Acoustical Society of America*, *79*(2), 500-507. doi:10.1121/1.393537.

Dahan, D., Magnuson, J., Tanenhaus, M., & Hogan, E. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes, 16*(5-6), 507-534. doi:10.1080/01690960143000074.

Damper, R., & Harnad, S. (2000). Neural network models of categorical perception. *Perception & Psychophysics*, *62*(4), 843-867. doi:10.3758/bf03206927.

Dar, M. (2016). *Effects of development on cross-language speech perception* (Unpublished doctoral dissertation). University of York, York, England.

Docherty, G. (1992). *The timing of voicing in British English obstruents*. Berlin: Foris Publications.

Eimas, P., Siqueland, E., Jusczyk, P., & Vigorito, J. (1971). Speech Perception in Infants. *Science, 171*(3968), 303-306. doi:10.1126/science.171.3968.303.

Grieser, D., & Kuhl, P. (1989). Categorization of speech by infants: Support for speech-sound prototypes. *Developmental Psychology*, *25*(4), 577-588. doi:10.1037//0012-1649.25.4.577.

Guenther, F. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, *102*(3), 594-621. doi:10.1037//0033-295x.102.3.594.

Harris, K., Bell-Berti, F., & Raphael, L. (1995). *Producing speech*. New York, NY: AIP Press.

Hattori, K. & Iverson, P. (2009). English /r/-/l/ category assimilation by Japanese adults: Individual differences and the link to identification accuracy. *Journal of the Acoustical Society of America*, *125*(1), 469-479.

Howell, P. (1983). The extent of coarticulatory effects: Implications for models of speech recognition. *Speech Communication*, *2*(2-3), 159-163. doi:10.1016/0167-6393(83)90017-1.

Iverson, P., & Kuhl, P. (1995). Mapping the perceptual magnet effect for speech using signal detection theory and multidimensional scaling. *Journal of the Acoustical Society of America*, *97*(1), 553-562. doi:10.1121/1.412280.

Keating, P. (2005). dprime analysis. [online] Phonetics.linguistics.ucla.edu. Available at: http://phonetics.linguistics.ucla.edu/facilities/statistics/dprime.htm [Accessed 21 Sep. 2017].

Kleber, F., Harrington, J., & Reubold, U. (2011). The Relationship between the perception and production of coarticulation during a sound change in progress. *Language and Speech*, *55*(3), 383-405. doi:10.1177/0023830911422194.

Kuhl, P. (1986). Reflections on infants' perception and representation of speech. In J. Perkell & D. Klatt (Eds.), *Invariance and Variability in Speech Processes* (pp. 19-30). Noorwood: Lawrence Erlbaum Associates.

Kuhl, P. (1991). Human adults and human infants show a “perceptual magnet effect” for the prototypes of speech categories, monkeys do not. *Perception & Psychophysics*, *50*(2), 93-107. doi:10.3758/bf03212211.

Kuhl, P. (2004). Early language acquisition: cracking the speech code. *Nature Reviews Neuroscience*, *5*(11), 831-843. doi:10.1038/nrn1533.

Kuhl, P., Conboy, B., 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*(1493), 979-1000. doi:10.1098/rstb.2007.2154.

Kuhl, P., 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*(2), F13-F21. doi:10.1111/j.1467-7687.2006.00468.x.

Kuhl, P., Tsao, F., & Liu, H. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*(15), 9096-9101. doi:10.1073/pnas.1532872100.

Kuhl, P., Williams, K., Lacerda, F., Stevens, K., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, *255*(5044), 606-608. doi:10.1126/science.1736364.

Livingston, K., Andrews, J., & Harnad, S. (1998). Categorical perception effects induced by category learning. *Journal of Experimental Psychology*: *Learning, Memory, and Cognition*, *24*(3), 732-753. doi:10.1037//0278-7393.24.3.732.

MacKain, K. (1982). Assessing the role of experience on infants' speech discrimination. *J. Child Lang*., *9*(03). doi:10.1017/s030500090000489x.

Maye, J., Weiss, D., & Aslin, R. (2008). Statistical phonetic learning in infants: facilitation and feature generalization. *Developmental Science*, *11*(1), 122-134. doi:10.1111/j.1467-7687.2007.00653.x.

Maye, J., Werker, J., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101-B111. doi:10.1016/s0010-0277(01)00157-3.

McMurray, B., & Aslin, R. (2005). Infants are sensitive to within-category variation in speech perception. *Cognition*, *95*(2), B15-B26. doi:10.1016/j.cognition.2004.07.005.

McMurray, B., Tanenhaus, M., & Aslin, R. (2002). Gradient effects of within-category phonetic variation on lexical access. *Cognition*, *86*(2), B33-B42. doi:10.1016/s0010-0277(02)00157-9.

Medin, D., & Barsalou, L. (1987). Categorization processes and categorical perception. In S. Harnad (Ed.), *Categorical Perception* (pp. 445-490). New York: Cambridge University Press.

Miller, J. (1994). On the internal structure of phonetic categories: a progress report. *Cognition*, *50*(1-3), 271-285. doi:10.1016/0010-0277(94)90031-0.

Miller, J. (1997). Internal structure of phonetic categories. *Language and Cognitive Processes*, *12*(5-6), 865-870. doi:10.1080/016909697386754.

Nam, Y., & Polka, L. (2016). The phonetic landscape in infant consonant perception is an uneven terrain. *Cognition*, *155*, 57-66. doi: 10.1016/j.cognition.2016.06.005.

Nartey, J. (1984). Coarticulation effects on fricative consonants across languages. *Journal of the Acoustical Society of America*,*75*(S1), S66. doi:10.1121/1.2021550.

Nittrouer, S., & Studdert-Kennedy, M. (1987). The role of coarticulatory effects in the perception of fricatives by children and adults. *Journal of Speech and Hearing Research*, *30*(3), 319-329. doi:10.1044/jshr.3003.319.

Pegg, J., Werker, J., & McLeod, P. (1992). Preference for infant-directed over adult-directed speech: Evidence from 7-week-old infants. *Infant Behavior and Development*, *15*(3), 325-345. doi:10.1016/0163-6383(92)80003-d.

Pierrehumbert, J. (2003). Phonetic Diversity, Statistical Learning, and Acquisition of Phonology. *Language and Speech*, *46*(2-3), 115-154. http://dx.doi.org/10.1177/00238309030460020501.

Pisoni, D., & Lazarus, J. (1974). Categorical and noncategorical modes of speech perception along the voicing continuum. *Journal of the Acoustical Society of America*, *55*(2), 328-333. doi:10.1121/1.1914506.

Pisoni, D., & Tash, J. (1974). Reaction times to comparisons within and across phonetic categories. *Perception & Psychophysics*, *15*(2), 285-290. doi:10.3758/bf03213946.

Polka, L., & Bohn, O. (1996). A cross-language comparison of vowel perception in English-learning and German-learning infants. *Journal of the Acoustical Society of America*, *100*(1), 577-592. doi:10.1121/1.415884.

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*(5), 2190-2201.

Polka, L., & Bohn, O. (2011). Natural Referent Vowel (NRV) framework: An emerging view of early phonetic development. *Journal of Phonetics*, *39*(4), 467-478. doi:10.1016/j.wocn.2010.08.007.

Polka, L., & Werker, J. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(2), 421-435. doi:10.1037//0096-1523.20.2.421.

Repp, B., & Crowder, R. (1990). Stimulus order effects in vowel discrimination. *The Journal Of The Acoustical Society Of America*, *88*(5), 2080-2090. http://dx.doi.org/10.1121/1.400105.

Repp, B., Healy, A., & Crowder, R. (1979). Categories and context in the perception of isolated steady-state vowels. *Journal of Experimental Psychology: Human Perception and Performance, 5*(1), 129-145. doi:10.1037//0096-1523.5.1.129.

Rivera-Gaxiola, M., Klarman, L., Garcia-Sierra, A., & Kuhl, P. (2005a). Neural patterns to speech and vocabulary growth in American infants. *Neuroreport*, *16*(5), 495-498. doi:10.1097/00001756-200504040-00015.

Rivera-Gaxiola, M., Silva-Pereyra, J., & Kuhl, P. (2005b). Brain potentials to native and non-native speech contrasts in 7- and 11-month-old American infants. *Developmental Science, 8*(2), 162-172. doi:10.1111/j.1467-7687.2005.00403.x.

Rosch, E. (1975). Cognitive reference points. *Cognitive Psychology, 7*(4), 532-547. doi:10.1016/0010-0285(75)90021-3.

Segal, O., Hejli-Assi, S., & Kishon-Rabin, L. (2016). The effect of listening experience on the discrimination of /ba/ and /pa/ in Hebrew-learning and Arabic-learning infants. *Infant Behavior and Development*, *42*, 86-99. http://dx.doi.org/10.1016/j.infbeh.2015.10.002.

Stevens, K., & Keyser, S. (2010). Quantal theory, enhancement and overlap. *Journal Of Phonetics*, *38*(1), 10-19. http://dx.doi.org/10.1016/j.wocn.2008.10.004.

Suzuki, N., Kitamura, T., Masaki, S., & Michi, K. (2001). MRI motion imaging of articulatory movement-coarticulation effects in velar consonants. *Journal of the Acoustical Society of America*, 110(5), 2656-2657. doi:10.1121/1.4777044.

Taylor, J.. (2008). Prototypes in cognitive linguistics. In P. Robinson & N. C.Ellis (Eds.), *Handbook of Cognitive linguistics & second language acquisition* (pp. 39-65). New York: Routledge.

Tsao, F., Liu, H., & Kuhl, P. (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*(4), 2285-2294. doi:10.1121/1.23382.90

Tsao, F., Liu, H., Kuhl, P. K., & Tseng, C. (2000). Perceptual discrimination of a Mandarin fricative-affricate contrast by English-learning and Mandarin-learning infants. In *Poster presented at the meeting of the International Society on Infant Studies*. England: Brighton.

Tsao, F., Liu, H., Kuhl, P., & Tseng, C. (2000). Perceptual discrimination of a Mandarin fricative-affricate contrast by English-learning and Mandarin-learning infants. In *Poster presented at the meeting of the International Society on Infant Studies*. Brighton: England.

Tsuji, S., & Cristia, A. (2013). Perceptual attunement in vowels: A meta-analysis. *Developmental Psychobiology*, *56*(2), 179-191. http://dx.doi.org/10.1002/dev.21179.

Tsuji, S., Mazuka, R., Cristia, A., & Fikkert, P. (2015). Even at 4 months, a labial is a good enough coronal, but not vice versa. *Cognition*, *134*, 252-256. http://dx.doi.org/10.1016/j.cognition.2014.10.009.

Tsushima, T., Shiraki, S., Yoshida, K., & Sasaki, M. (2003). On stimulus order effects in discrimination of nonnative consonant contrasts. *Acoustical Science And Technology*, *24*(6), 410-412. http://dx.doi.org/10.1250/ast.24.410

Tsushima, T. (2007). Asymmetries in perception of an American English /r-l/ by adult Japanese learners of English. *Journal of the Japan Society for Speech Sciences*, 8, 45-62.

Tsushima, T. (2011). Effects of stimulus order in categorial AX discrimination training on improvements of the ability to perceive the English/b/-/v/contrast among Japanese learners of English. *The Journal Of Communication Studies*, (33), 267-292.

Tsushima, T., Shiraki, S., Yoshida, K., & Sasaki, M. (2003). On stimulus order effects in discrimination of nonnative consonant contrasts. *Acoustical Science And Technology*, *24*(6), 410-412. http://dx.doi.org/10.1250/ast.24.410.

Tsushima, T., Shiraki, S., Yoshida, K., and Sasaki, M. (2005). *Stimulus order effects in discrimination of a nonnative consonant contrast, English /b-v/, by Japanese listeners in the AX discrimination procedure*. Unpublished paper presented at 'First Acoustical Society of America Workshop on L2 Speech Learning', Vancouver, Canada.

Tversky, A., & Gati, I. (1978). Studies of similarity. In E. Rosch & B. Lloyd (Eds.), *Cognition and Categorization* (pp. 79-98). Hillsdale: Lawrence Erlbaum.

Vihman, M. (1996). *Phonological development*. Cambridge: Blackwell.

Werker, J., Gilbert, J., Humphrey, K., & Tees, R. (1981). Developmental aspects of cross-language speech perception. *Child Development*, *52*(1), 349-355. doi:10.2307/1129249.

Werker, J., & Tees, R. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology/Revue Canadienne De Psychologie*, *37*(2), 278-286. doi:10.1037/h0080725.

Werker, J., & Tees, R. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *7*(1), 49-63. doi:10.1016/s0163-6383(84)80022-3.

Werker, J., & Tees, R. (2002). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, *25*(1), 121-133. doi:10.1016/s0163-6383(02)00093-0.

Wanrooij, K., Boersma, P., & Zuijen, T. (2014). Fast phonetic learning occurs already in 2-to-3-month old infants: an ERP study. *Frontiers In Psychology*, *5*, 77. http://dx.doi.org/10.3389/fpsyg.2014.00077.

Yeung, H., & Werker, J. (2009). Learning words' sounds before learning how words sound: 9-Month-olds use distinct objects as cues to categorize speech information. *Cognition*, *113*(2), 234-243. doi:10.1016/j.cognition.2009.08.010.

Yoshida, K., Pons, F., Maye, J., & Werker, J. (2010). Distributional phonetic learning at 10 months of age. *Infancy*, *15*(4), 420-433. doi:10.1111/j.1532-7078.2009.00024.x.

Appendix: List of words used in Experiment 3

Unaspirated	Aspirated
tʃa:l	tʃʰa:l
tʃa:p	tʃʰa:p
tʃak	tʃʰak
tʃal	tʃʰal
tʃoti	tʃʰoti
tʃiɽna	tʃʰiɽna
tʃi:n	tʃʰi:n
tʃoũn	tʃʰoũn
tʃour	tʃʰouɽ
tʃup	tʃʰup
tʃupkay	tʃʰupkay
tʃu:na	tʃʰu:na