Perception, production and perceptual learning in the second language: a study of perceptual learning by L1 Bengali speakers of L2 English

Jacqueline Ingham
University of Sheffield

Abstract

This study tests for evidence of perceptual speech learning with reference to two cross-linguistic perception models, the Perceptual Assimilation Model and the Speech Learning Model. A simulated longitudinal study is piloted with two adult native Bengali subjects with L2 English and differing L2 linguistic experience. The perception and production data of L2 English word-initial obstruents /p/ - /b/, /f/ - /v/ and /b/ - /v/, which are both shared and not shared in the L1, are compared and analysed for evidence of speech learning in intelligibility between the two learners. The context for this is whether perception-led classroom-based pronunciation training may improve adult L2 pronunciation of word-initial obstruents. Results show that the simulated longitudinal model may provide a window on perceptual learning. Evidence of learning in both perception and intelligibility in the production of word-initial obstruents /p/, /f/ and /v/ is detected in the participants in this study. It is argued, however, that whilst there is some evidence that perceptual speech learning may occur over time, further research is necessary to investigate speech learning at different stages of experience in the L2.

Key words: *Perceptual speech learning; Second Language Acquisition; Bengali speakers of English.*

Introduction

The debate on adult accented speech has developed over several decades (e.g. Flege, 1995; Lenneberg, 1969; Scovel, 2000) and, whilst the current focus may have shifted, age, input and ultimate attainment in second language acquisition remain highly topical (e.g. Montrul, 2010; Muñoz and Singleton, 2011; Rothman, 2008)[1]. Initially, several studies focussed on the similarities and differences between child and adult first (L1) and second (L2) language acquisition, arguing for a biologically timed critical period beyond which neural plasticity is atrophied, preventing the post-pubescent learner from achieving target-like L2 speech (e.g. Scovel, 1969; Oyama, 1976, Patowski, 1990; Long, 2007). Other studies rejected claims of a critical period, with evidence of adult L2 learners able to produce unaccented speech (Snow & Hoefnagel-Höhle, 1977; Bongaerts, 1999), and pre-pubescent children unable to produce unaccented L2 speech (Flege & Eefting, 1987)[2]. Studies such as that by Flege (1991) and Flege and Eefting (1987) argue that any advantage gained by learning to speak an L2 in the pre-pubescent years is tempered not by a critical period, but by the quality of L2 input. Proponents of perception based theories include Flege (1995), who proposes that L2 production is led by the perception and classification of new or similar L2 sounds through the L1 phonological system, and Best (1995) who claims new L2 sounds are subject to differing degrees of assimilation according to existing L1 categories.

Although teaching methodology is not the prime concern of this study, the critical period theories have arguably had extensive influence on the theory and practice of English as a foreign or second language (EFL, ESL) teaching methodology and language planning (Scovel, 2000). This is seen in the dominance of articulatory-based training for adult pronunciation practice (Rochet, 1995) and the replication of child L1 acquisition for older L2 learners (Scovel, 2000). The role of perception in promoting accurate production has arguably had little impact on L2 classroom practice (Rochet, 1995), despite a number of cross-linguistic perception studies (e.g. Best, 1995; Best, Halle, Bohn & Faber 2003; Best &

[1] The use of the term 'accented' and 'target-like' herein refers to how closely a native listener of the target language perceives and deems intelligible the segmental production of a second language speaker (see Higgins, 2003; Jenkins, 2005; Jenkins, 2006 for discussion on World Englishes and standard versions of English).

[2] There has been much debate regarding methodology, for instance in the analysis and representation of graphical data in critical period studies (e.g. Birdsong, 2005), and in areas such as the weight placed on accent compared to comprehensibility and intelligibility (e.g. Derwing and Munro, 1997) and the reliability of native speaker benchmarking (e.g. Bongaerts, 1999; Bongaerts, van Summeren, Planken, and Schils 1997; Rothman, 2008).

Tyler, 2007; Flege, 2003). It is this apparent lack of uptake of perception-based theories in the teaching of adult L2 pronunciation which motivates this current study. The question addressed is whether or not perceptual learning occurs over time, with the implications of this being whether perceptual learning could be harnessed in classroom-based pronunciation teaching.

There are a number of studies which test the perception and production of specific consonant or vowel phonemes by adult L2 learners of English or other languages, such as those in support of L2 perception based on L1 categories (e.g. Best, McRoberts & Sithole, 1988 on non-native Zulu click discrimination), or those which challenge the perception-production correlate (e.g. Chan [2014] on Cantonese ESL learners). Whilst other studies detail the effects of perceptual training on perception and production (e.g. Rochet, 1995; Hanulíková, Dediu, Fang, Bašnaková, & Huettig, 2012), there are comparatively fewer studies testing for evidence of perceptual learning of L2 sounds by adult L2 learners over time (e.g. Guion, Flege, Akahane-Yamada & Pruitt, 2000).

In this paper, perception and production data of word-initial obstruents from two L1 Bengali speakers of L2 English is examined for evidence of perceptual learning. An experimental simulated longitudinal test attempts to replicate the conventional longitudinal study by extrapolating between an initial state learner and an experienced or bilingual learner according to the predictions of two distinct but compatible perception-based models. It is proposed that this methodology allows insight into perceptual development over time, with evidence of perceptual learning and new category formation, as well as modification of similar categories in the L1 and L2. However, this perceptual learning is measured over a simulated, but significantly lengthy period of time, and the return in terms of improved pronunciation for such extensive exposure to quality L2 input may need to be considered in terms of applicability to classroom learning.

The following section discusses two cross-language perception models relevant for the current study. I describe the experimental methodology, and the predictions for perception and speech learning are detailed thereafter. A discussion of the results is given followed by the conclusion section.

Two models for predicting perception and production of L2 sounds

One of the problems in testing for evidence of perceptual learning is that it is proposed to be a lifelong faculty (Flege, 1995) and category formation and speech learning may involve several years of quality target-language input (Flege, 1995; Guion et al., 2000). Two cross-language perception models with which it is seemingly possible to examine the learnability of L2 consonant contrasts by L2 learners proficient at both ends of the learning spectrum, are the Perceptual Assimilation Model (PAM; Best et al., 1988; Best, 1995) and the Speech Learning Model (SLM; Flege, 1995). Whilst there are studies which test PAM against SLM (e.g. Rohena-Madrozo, [2013] on occluded voiced stops by L1 Spanish subjects), PAM and SLM may also be seen as complimentary models (Best & Tyler, 2007). Used in tandem, PAM and SLM may test for learnability of the perception of L2 sounds with respect to language experience (e.g. Guion et al., [2000] on the perception of English consonants by adult L1 Japanese speakers).

The Perceptual Assimilation Model (PAM)

PAM (e.g. Best et al., 1988; Best, 1995) proposes that the ability of naïve adult L2 learners to discriminate between non-native phonological contrasts is commensurate with how the features of contrasting L2 phones are assimilated and categorised to existing L1 phonological categories. This concerns how the articulatory properties, such as place and degree of constriction of an L2 sound are perceived in relation to the nearest L1 sound. The perceived distance between L1 and L2 sounds affects the learner's ability to discriminate between L2 contrasts. In practice, this means that some contrasting pairs of L2 sounds are proposed to be 'excellent' and easier to detect than 'poor' examples of an L1 category (Best, 1995).

PAM defines six assimilation patterns (Table 1) allowing predictions to be made regarding the discriminatory ability of the learner to detect the contrast between two L2 sounds. The ability to discriminate is rated according to whether the L2 phones are considered good or bad examples of the L1 category. In this respect, a two-category assimilation whereby two L2 phones are assimilated to two corresponding L1 categories is the most accurate in terms of discriminatory ability of the learner, whereas two L2 sounds assimilated to one L1 sound, as a single-category assimilation may cause poor discriminatory

ability. Not all L2 sounds are considered speech sounds. If an L2 segment is not perceived as a speech sound, then it is not assimilable within the L1 phonological space. An L2 speech sound which is within the L1 phonological space, but which does not correspond to any particular native category, is considered uncategorisable.

Table 1: The PAM assimilation patterns for non-native contrasts. Note: Adapted from Best (1995, p. 125).

Category	**Assimilation pattern** **X → assimilates to Y**	**Predicted discrimination**
Two-Category (TC Type)	Two L2 sounds → Two L1 sounds	Excellent
Single-Category (SC Type)	Two L2 sounds → One L1 sound	Poor
Category-Goodness (CG Type)	Two L2 sounds → One L1 sound. One L2 sound is a good example of the L1 sound, the other is a poor example	(Very) good (to moderate)

The Speech Learning Model (SLM)

Whilst PAM provides a framework for prediction of the discriminatory perceptual abilities between non-native L2 contrasts by inexperienced adult L2 learners, SLM focuses on the speech learning of very experienced learners and bilinguals, predicting how accurately experienced learners may both perceive and, importantly, produce L2 sounds with respect to the potential for lifelong learning of both perception and speech.

According to SLM, as proposed in the version set forward by Flege (1995), the L1 mechanisms used to create and store phonetic categories are available throughout adulthood and are applicable to individual L2 sounds, allowing new categories to be created for phonetically different L2 sounds when distant enough from the nearest category in the L1. Furthermore, bilinguals operate two language systems within the same phonological space, and significant effort is made to maintain the phonetic contrasts between the L1 and L2 categories. This is significant because according to this version of SLM, L1 'phonetic categories' are susceptible to the influence of the properties of L2 sounds.

Similar to PAM, SLM proposes that L2 learners perceive auditory sounds through the L1 phonological system, but in contrast to PAM, SLM claims that the greater the perceived distance between an L2 sound and the nearest L1 sound, the more likely it is that a new category will be formed. Furthermore, L2 speech will be more or less accented according to

how similar the representation of the new category is to that of a native speaker of the target language. In other words, quality and quantity of L2 exposure remain integral to the success of new category formation (Guion et al., 2000).

Alongside the constraints imposed by advancing age, however, the formation of a new category depends upon whether an L2 sound is perceived as 'new' or 'similar'. This is argued to be subject to equivalence classification, a process also at work in child L1 acquisition, which allows infants to identify a particular phonetic category even though a phone may be produced variably due to speaker idiosyncrasies or the surrounding phonetic environment (e.g. Flege, 1987, 1995). In adult and older child L2 language learning, Flege (1987) proposes that equivalence classification prevents discrimination between articulatory 'similar' segments, which are present in both the L1 and L2. This is illustrated in the case of L1 English speakers learning L2 French, whereby learners will identify that /t/ is a 'similar' phone, found in both the L1 and L2 (Flege, 1987). However, whilst 'similar', the French and English /t/ are not 'identical', (Flege, 1987) with differences in both Voice Onset Time and place of articulation (English /t/ is long-lag stop with alveolar place of articulation and French /t/ is a short-lag stop with dental place of articulation). Flege (1987) claims that equivalence classification prevents the learner from making a new phonetic category for the 'similar' L2 phone, and that target-like L2 production is subsequently inhibited, which may over time even cause amalgamation of the L1 and L2 qualities to a single category.

On the other hand, equivalence classification does not interfere with the perception and category formation of new L2 phones, which are acoustically distinct from those phones present in the L1, such as the 'new' L2 French /y/ for L1 English learners of French (Flege, 1987). Whilst Flege (1987) suggests that the French /y/ might initially be identified as /u/ by L1 English speakers, it is proposed that speech learning will occur so that highly experienced L1 English speakers of L2 French will produce target-like French /y/. The principles of speech learning (SLM; Flege, 1995) are set out in the following list adapted from Flege (1995, p. 239):

- L1 and L2 sounds are perceptually related at an allophonic level.

∠

- If a bilingual detects phonetic differences between L2 and closest L1 sound, a new phonetic category can be created.

∠

- Phonetic differences will be detected if there is a greater perceived variance between the L2 and closest L1 sound.

∠

- As age of learning increases the probability of detecting phonetic differences between L1 and L2 sounds, or L2 sounds which are not contrasted in the L1, diminishes.

∠

- Equivalence classification may obstruct the creation of a new L2 category so that perceptually linked L1 and L2 sounds will be processed into a single category, influencing production of both L1 and L2 sounds to sound the same.

∠

- A bilingual's category for an L1 sound may be different from that of a monolingual speaker if the L2 category is pushed away from an existing L1 category to preserve contrast.
- Sound production eventually matches with phonetic category representation.

Perception and production

According to both these models, the ability to discriminate between L2 contrasting consonant sounds depends on the how the phonetic features of the L2 sound is perceived in relation to those of existing L1 sounds. The perceived distance between the L1 and L2 sound determines how accurately the L2 sound may be assimilated and categorized in relation to that of a native speaker of the target language. Whilst PAM does not make predictions on L2 production, perception precedes production in terms of the initial assimilation of articulatory gestures in L1 acquisition, which defines the discriminable phonetic distinctions underpinning L1 phonological contrasts against which non-native segments are perceived (Best, 1995). In terms of SLM, accurate perceptual L2 tokens are prerequisite to promote the sensorimotor learning in the production of target-like speech sounds (Flege, 1995).

The present study

In the present study, two L1 Bengali speakers with different experience of L2 English were tested for evidence of learning in both perception and production of word-initial consonants. The simulated longitudinal study in this experiment was designed as an initial pilot to test whether evidence of perceptual learning can be extrapolated between learners from the same L1 background with differing L2 linguistic experience, using the predictions of PAM and SLM at both the initial and advanced/bilingual stages of learning respectively. L1 Bengali speakers are relatively underrepresented in cross-language perception studies.

Participants

Two L1 speakers of Bengali with L2 English participated in this initial pilot study. The participants were selected following a cross-linguistic pre-pilot study in conjunction with a contrastive analysis of word-initial consonant phonemes shared and not shared in the L1 and L2 (see Table 4). A learner background summary is presented in Table 2. Levels of L2 English were evaluated by means of self-assessment, and although the L2 speaking level was not validated during this study, self-reports of target-like pronunciation in the L2 have similarly been recorded in other studies, such as that by Flege, Munro and MacKay (1995) with native Italian speaking subjects.

Table 2: Learner background of participants

	Participant A	**Participant B**
Self-assessed level in speaking L2 English	Elementary	Advanced
and % of L2 usage per week	20%	80%
Age of arrival in UK	31	7 - 10 Settled permanently in the UK at age 11
Age at testing	47	35

A point of discussion is age, age of arrival and age of testing of the participants. Flege (1987) distinguishes between young children as one category of learner and older children and adults as another, but it is not clear at which point a child progresses from being young to old. The age of arrival (AOA) in the UK of Participant B is particularly relevant, especially as permanent residency and full-time education in L2 English did not occur until the participant was aged eleven. Flege (1995) claims that the impact of AOA on the perception and production of L2 sounds that are not shared in the L1 remains unclear. Studies in support of a critical period also propose different critical ages for L2 speech learning. Oyama (1976) identifies a sensitive-period for acquiring an L2 phonological system with an AOA of twelve, regardless of the length of stay, whereas Asher and García (1969), propose children with an AOA of between one and six years old proved more likely to acquire target-like speech than those with an older AOA. The situation is much the same in perception-led L2 speech studies (see Flege for a brief review of some of the studies on AOA, 1995), although both camps agree that the earlier the age of L2 learning, the better for L2 pronunciation.

Participant B started learning English as an older child or adult, and has subsequently had twenty-four years of quality L2 input. Importantly, as this study is concerned with evidence of speech learning by comparing data from the inexperienced and experienced L2 learner, the L2 input has occurred during the maturational and early adult years, where the learner is potentially more receptive to perceptual learning and new phonetic category formation than in the later or advanced years of learning (Flege, 1995). Interestingly, Participant B did not select 'bilingual' as an option to self-describe L1 and L2 usage, and the terms 'advanced' and 'experienced' are used with consideration to the self-assessment.

Although Participant A had resided in the UK for some sixteen years at the time of testing, exposure to the L2 has been extremely limited and Participant A is considered an inexperienced learner, despite the length of stay in the UK. The relevance of this to the present study is that it is assumed that the predictions of PAM for the discriminatory ability of the inexperienced adult L2 learner in the perception of word-initial L2 consonant contrasts are applicable, and will subsequently provide the baseline for testing for evidence of speech learning in comparison to the experienced participant's data.

Stimuli

Five L2 English obstruents were tested in three contrasting minimal pairs in L2 English: /p/ - /b/, /f/ - /v/, and /b/ - /v/. The tested phonemes are set out in Table 3 in English and Bengali (Ferguson and Chowdhury, 1960). The L2 English /p/ and /b/ phonemes are considered shared in the L1, although Bengali has a four-way contrast with aspirated and unaspirated contrasts as well as voiced and voiceless counterparts. The English /f/ and /v/ contrast is not shared in the L1, and the /b/ - /v/ contrast has both a shared and not shared phoneme between the L2 and the L1.

The target phonemes were tested in word-initial position of monosyllabic CVC words, such as 'pin' - 'bin'.

Table 3: Comparison of tested phonemes in English and Bengali

			Bilabial	**labiodental**
plosive	Bengali	voiceless	p p^h	
		voiced	b b^h	
	English	voiceless	P	
		voiced	B	
fricative	Bengali	voiceless		
		voiced		
	English	voiceless		f
		voiced		v

Perception test materials, procedure and method of analysis

The stimulus material was designed to test the subjects' perception of contrasting word-initial obstruents in an AX word discrimination test. These tests have been used to test perception of non-native contrasts in a number of studies, including cross-linguistic studies (e.g. Best et al., 1988). The test is relatively straightforward and requires the individual to identify whether two phonetic tokens are the same (X is identical to A) or in some way different (X is not identical to A). An AX discrimination test was chosen in preference to and AXB type test because it requires less strain on the memory (Strange and Shafer, 2008), and is arguably more appropriate for inexperienced learners.

The stimuli consisted of a pre-recorded set of seventeen pairs of CVC monosyllabic words, with additional distractors and practice examples, in an approximate ratio of 2:1 for contrasting sounds over same-sound minimal pairs. The material contained only real words and only word-initial sounds were tested which were:

i) shared in the L1 and L2 /p/ - /b/
ii) not shared in the L1 and L2 /f/ - /v/
iii) /b/ - /v/ one token shared and one token not shared in the L1 and L2.

The recorded perception test was presented in the following format:

Pre-recorded female English native speaker says 'number one' (delay 2 seconds). Male English native speaker says 'sip (delay 4 seconds) ship (delay 4 seconds). Female speaker says 'number two' (delay 2 seconds). Male speaker says 'pin (delay 4 seconds) 'pin'.

Answer sheet: Participants tick one of two columns 'Same' or 'Different'.

The inexperienced and experienced L2 English learners completed the perception test during one sitting. The data were analysed according to accurately perceived L2 contrasts, which were then calculated into a percentage.

Production test materials, procedure and method of analysis

The production test stimuli included the tokens from the perception test, presented in a random order as a list of words with distractors and examples to a total of 25 tokens. The test was delivered as an imitation or repetition procedure, and the learners were asked to repeat an auditory prompt of pre-recorded words, a procedure Bradlow, Pisoni, Arkahane-Yamada & Tohkura (1997) have used with both visual and auditory prompts. Visual prompts were not included in this test in an attempt to reduce deliberateness of speech. Similarly, only a short delay was given after the audio prompt in order to force a quick and unstudied response from the participants. The recording was put onto a personal sound system, and played through headphones. Subjects followed a 'listen and repeat' sequence and the production data was recorded onto a laptop. The procedure was as follows:

Pre-recorded male English native speaker says 'number one' (delay 1½ seconds). The male speaker says 'ship' (delay 10 seconds).
Production: The subject listens to the recording and repeats the word within the 10-second pause (utterance recorded).

The recorded speech production of the two participants was subsequently played to three adult native speakers of English: a primary school teacher, a university student and a secondary school teacher, who were asked to transcribe the words in an intelligibility judgement test, as exemplified in Table 4.

Table 4: Imitation production test - sample analysis of response data

Production Participant A	**Transcriber 1**	**Transcriber 2**	**Transcriber 3**	**Analysis**
fat	pan	an	Van	0

Production Participant B	**Transcriber 1**	**Transcriber 2**	**Transcriber 3**	**Analysis**
fat	fat	fat	Fat	1

Corroborative analysis of the correct representation of the target phoneme by all three transcribers was required in order to evidence accurate production of a phoneme. A percentage of accurately produced L2 target segments was then calculated as a production accuracy score for each contrasting segment, and a 70% accuracy rate was considered evidence of category formation[3].

Predictions for evidence of speech learning

Predictions for the perceptual ability of the inexperienced learner

Table 5: Participant A - Predictions for perceptual discriminatory ability

	PAM TC Type: Two L2 sounds → Two L1 sounds = excellent discrimination
1. /p/ - /b/ contrast	L2 English /p/ assimilates to L1 Bengali /p/ L2 English /b/ assimilates to L1 Bengali /b/ **= excellent discriminatory ability**
	PAM UU Type: Two L2 sounds → No L1 sound = poor to excellent discrimination
2. /f/ - /v/ contrast	L2 English /f/ assimilates as poor example of L1 Bengali /pʰ/ or /p/ L2 English /v/ assimilates as poor example of L1 Bengali /bʰ/ or /b/ **= good discriminatory ability.** The L2 sounds are both assimilated within the L1 phonological space, but are uncategorisable within L1 categories.
3. /b/ - /v/ contrast	PAM UC Type: Two L2 sounds → a) one L2 sound to one L1 sound category, and b) one L2 sound uncategorised = very good discrimination L2 English /b/ assimilates to L1 Bengali /b/ L2 English /v/ assimilates to L1 Bengali /b/ or /bʰ/ PAM = very good discriminatory ability Problem with uncategorised sounds, which closely resemble categorised sound in UC Type category. **Revised prediction = poor discriminatory ability**

The predictions for the perception of the target sounds by the inexperienced learner are made according to the PAM framework and are set out in Table 5. It is expected that the learner

[3] A 70% criterion level for category formation was adopted for this study. See Lakshmanan and Selinker (2001) for a discussion on criterion levels in morpheme and feature acquisition.

will show excellent ability to discriminate between L2 /p/ - /b/, and good and poor discriminatory ability for /f/ - /v/ and /b/ - /v/ contrasts respectively.

The prediction for L2 English /f/ - /v/ contrast is predicted as good for the inexperienced learner because both L2 sounds are uncategorized, but they are equally phonetically poor examples relative to the nearest L1 categories of /pʰ/ or /p/ and /bʰ/ or /b/. However, although according to PAM, the learner should have very good discriminatory ability discerning between L2 English /b/ - /v/ contrast, this prediction is revised to poor discriminatory ability, following the study by Guion et al., (2000). In this study (Guion et al., 2000 p. 2721), the authors identify the potential for modification to PAM regarding the uncategorized - categorized UC Type category when the uncategorized sound (e.g. L2 English /v/ to L1 /b/ or /bʰ/), is in close proximity to the categorized sound (e.g. L2 English /b/ to L1 /b/) within the phonological space. With this in mind, the inexperienced L1 Bengali learner of L2 English is predicted to have poor rather than excellent discriminatory ability of L2 /b/ - /v/ contrast.

Predictions for evidence of speech learning

Table 6: Predictions for Participant B and predictions for speech learning

L2 sounds	Predictions for Participant B	Predictions for evidence of speech learning between Participant A & B
1. L2 /p/ - /b/	L2 /p/ and L1 /p/ = good match L2 /b/ and L1 /b/ = good match	**No evidence of learning predicted** between subject A and B in perception or production, as both subjects should be able to perceive and produce the L2 contrast as also present in the L1.
2. L2 /f/ - /v/	New category formation is predicted to have occurred for L2 /f/ and L2 /v/ as 'new' sounds	**Evidence of learning predicted** between subject A and B, as subject B predicted to both perceive and produce L2 /f/ and /v/ more accurately and consistently than subject A.
3. L2 /b/ - /v/	New /v/ category and L2 /b/ and L1 /b/ = good match	**Evidence of learning predicted** between subject A and B, as subject B is predicted to outperform subject A in perception and production of L2 /b/ and /v/ contrast.

The predictions for the perception and production of the target sounds by the experienced learner are made according to the hypotheses of SLM, and in comparison to the expected discriminatory ability of the inexperienced learner. These are set out in Table 6. Evidence of

learning is predicted to occur in L2 consonants /f/ and /v/ which are not shared in the L1, and the experienced learner is expected to have created new categories for /f/ and /v/ phonemes, which is predicted to be reflected in significantly higher levels of accuracy in perception and production than the inexperienced learner.

Results and discussion

A combined set of results for both participants is set out in Table 7. Figure 1 illustrates the perception test, and Figure 2 the production test.

Table 7: Perception and production data

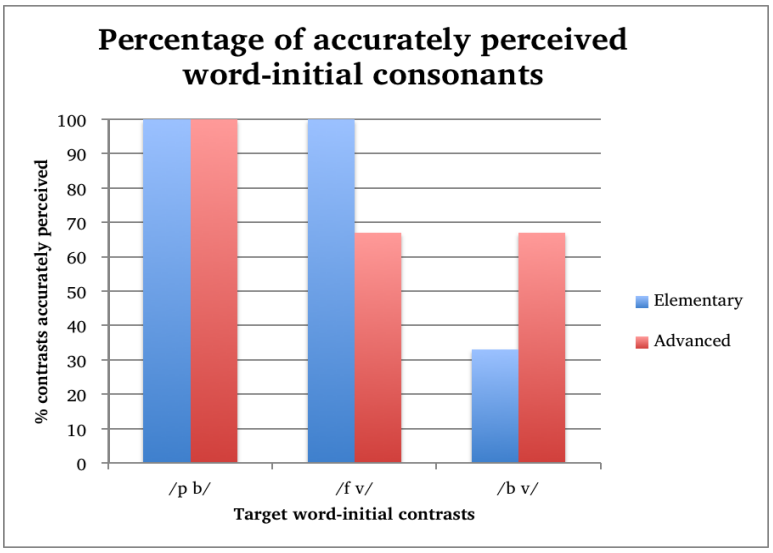
	Phoneme	Participant A %	Participant B %
Perception	/p/ - /b/	100	100
	/f / - /v/	100	67
	/b/ - /v/	33	67
Production	/p/	40	80
	/b/	82	74
	/f/	50	83
	/v/	0	59

Both participants achieved 100% accuracy in discriminating between L2 English /p/ - /b/ contrast, as predicted according to both PAM and SLM. The ability of the inexperienced learner to discriminate between L2 English /f/ - /v/ contrast is higher than predicted, again achieving 100% accuracy. PAM predicts poor to excellent discriminatory ability for UU Type category assimilation, and the result here perhaps indicates that the perceived distance between the uncategorized L2 sounds and the L1 categories of the inexperienced learner was significantly less than anticipated in this study. It could be that the voicing distinction between L2 /f/ and /v/ was a sufficient phonetic distinction to promote excellent discriminatory ability, and that this result should therefore be considered commensurate with PAM UU Type assimilation predictions.

The inexperienced learner's perception of the L2 English /b/ - /v/ voiced - voiced contrast, where the learner shows evidence of lower perceptual ability, is in line with the

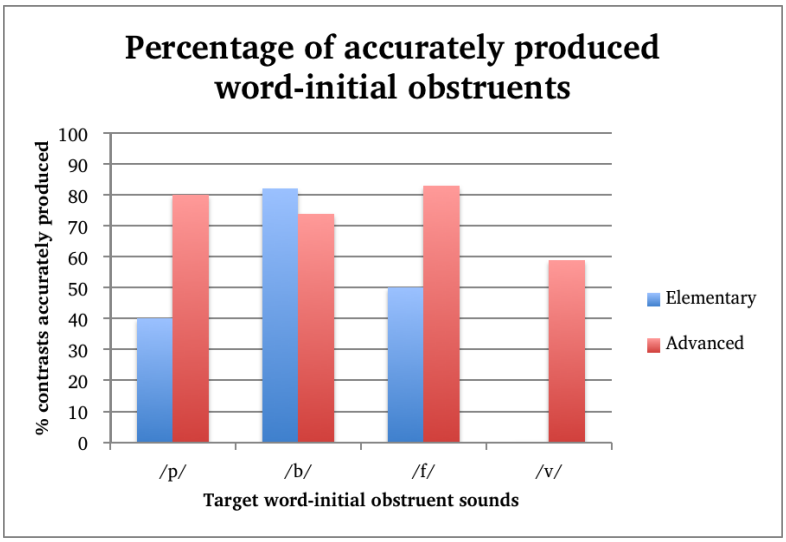
revised PAM prediction as discussed in Section 4.1. The poor discrimination of L2 /b/ - /v/ not only reflects the perceptual problems encountered when an uncategorized sound is in close proximity to an L2 sound which has been categorized within an L1 phonetic category, but also the effect phonetic similarity of two contrasting L2 phones (such as voicing) can have on the ability for inexperienced learners to discriminate between L2 contrasts which are not present in the L1. Interestingly, in evidence of perception preceding production, the inexperienced learner shows good production accuracy in producing L2 /f/, but is unable to produce target-like /v/ in any context (Figure 2). Furthermore, although the inexperienced learner's production score for evidence of /v/ was 0%, all three transcribers noted /b/ for every single instance that the inexperienced learner attempted to produce word-initial /v/.

Figure 1: Percentage of accurately perceived word-initial obstruents



It would appear that the predictions based on PAM provide an accurate indication of the inexperienced L2 English learner's ability at the initial stage of L2 learning to discriminate between contrasting L2 consonants /p/ - /b/, /f/ - /v/ and /b/ - /v/, which are shared and not shared in L1 Bengali, and provide a reliable baseline from which evidence of perceptual learning can be measured. Consistent with the predictions of SLM, analysis for evidence of speech learning between the experienced and inexperienced learner combines the perception and production data (Table 7), and the production results as presented separately in Figure 2.

Figure 2: Percentage of accurately produced word-initial obstruents



Focusing on the production of the two consonants /f/ and /v/, which are not shared between the L1 and L2, and assuming that a 70% accuracy in production is commensurate with category formation, it would appear that the experienced learner has created a new category for /f/, and is approaching a new category /v/. As the experienced learner has perception accuracy rates of 67% for both /f/ - /v/ and /b/ - /v/ contrasts, there is some but not total consistency in perception and production, and according to SLM, there is no reason to suspect that for this subject these categories will not continue to improve in accuracy with continued speech learning.

This is in sharp contrast to the data for Participant A, especially in terms of the lack of /v/ production in contrast to the comparative accuracy in the production of /f/ and the perception of the /f/ - /v/ contrast. The difference between Participant A and Participant B in production and perception of these sounds may be somewhat obscured by these comparative results. However, speech learning may not be linear and extrapolating between learners with yet different L2 experience may illuminate the development of new category formation. Still, there is no clear explanation as to why in production a new category for L2 /f/ would be more accurate than a new category for /v/. The experienced learner shows equal ability in perception of /f/ and /v/, and both categories are new and not 'similar' to L1 phonetic categories. The apparent emergence of /f/ before /v/ in category formation is perhaps not only relative to the two Bengali participants in this study (e.g. Jehma and Phoocharoensil [2014] cite similar evidence from Pattani-Malay learners of L2 English). It is possible that the experienced learner

shows evidence of bilingual merging (Flege, 1987), whereby the L1 and L2 sounds are perceptually combined into a single category, influencing perception and production of the L1 as well as the L2 when compared to that of a monolingual speaker. This requires further testing of the L1 alongside the L2, but merging would provide explanation for the lower accuracy rate of the experienced learner's perceptual data compared to the inexperienced learner for /f/ - /v/, and the comparable score for the /f/ - /v/ and /b/ - /v/ contrast (both 67% accuracy).

Much has been written on the acquisition of voiceless stops (Flege, 1995), and it is interesting to note that in the context of this study, Participant A has a low production accuracy rate for L2 /p/ compared to L2 /b/ (Figure 2), even though these phonemes are shared with the L1, and perception of /p/ - /b/ contrast is excellent. It should be remembered, however, that PAM does not make predictions regarding production. The higher accuracy rate of the experienced learner in the production of L2 voiceless stop /p/ compared to that of the inexperienced learner, is proposed to be evidence of speech learning, which was not anticipated in this study as both phonemes have counterparts shared in the L1. However, as noted in Table 3, Bengali has a four-way distinction between /p/, /pʰ/, /b/ and /bʰ/ compared to the two-way voiced - voiceless distinction in English /p/ - /b/. Ferguson and Chowdhury (1960) note that Bengali /pʰ/ and /bʰ/ are produced either as an aspirate; a stop followed by an aspirated release, or as a spirant, such as a fricative. The higher accuracy rates for the experienced learner in the production of L2 /p/ may arguably be evidence of quite fine phonetic speech learning taking place within a similar L1/L2 sound (Flege, 1987), which if tested against the /p/ sounds of an L1 Bengali monolingual speaker and learners with differing levels of L2 experience, may provide greater insight into merging and speech learning.

Conclusion

This experimental pilot study has arguably shown evidence of speech learning between an experienced and inexperienced native Bengali speaker in the perception and production of L2 English word-initial consonants. That is particularly L2 English voiceless stop /p/, which had not been predicted, and both voiceless and voiced fricatives not shared in the L1, /f/ and /v/. Further phonetic analysis of L2 production data in conjunction with L1 production of

nearest related phonemes might help to elucidate category formation and merging by the experienced learner in this study.

As an experimental design, this pilot aimed to test whether a simulated longitudinal study, drawing data from an adult inexperienced in the L2 and comparing it with a highly experienced L2 speaker from the same L1 background, could be used to ascertain whether L2 speech learning could occur over time by a process of extrapolation. The limitations of this study must be taken into account, especially the very small sample size of only two participants and the limited number of phonemes tested. However, it is tentatively suggested that the findings from this experimental pilot show that evidence of speech learning may potentially be extrapolated between learners with different L2 experience, but this must extend to larger subject groups with differing levels of experience to determine whether the development of speech learning accommodates a non-linear path of category formation and merging.

Regarding whether or not there should be a reconsideration of current classroom-based pronunciation teaching to include perception and speech learning alongside articulatory practice, this study can only contribute in terms of suggesting that more simulated longitudinal studies (with modification in light of the limitations from this study) are carried out to ascertain perceptual learning at different intervals of L2 development with differing L1 groups and phonemes. The evidence of speech learning in this experiment was after twenty-four years of naturalistic quality L2 input, and new category creation was not deemed complete. The application of perceptual learning in the classroom may be more feasible if patterns of speech learning can be extracted from differing stages of experience in the L2.

References

Asher, J. J., & García, R. (1969). The optimal age to learn a foreign language. *Modern Language Journal*, 53 (5), 334-341.

Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171-206). Timonium, MD: York Press.

Best, C. T., Halle, P., Bohn, O. S., & Faber, A. (2003). Cross-language perception of nonnative vowels: Phonological and phonetic effects of listeners' native vowels. In M. J. Solé and

J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, (pp. 2889-2892). Barcelona, Spain: Univesitat Autònoma de Barcelona.

Best, C. T., McRoberts, G. W., & Sithole, N. M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English-speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance*, 14 (3), 345-360.

Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception. Commonalities and complementarities. In M. J. Munro and O. S. Bohn (Eds.), *Second language speech learning: The role of language experience in speech perception and production* (pp. 13-34). Amsterdam: John Benjamins.

Birdsong, D. (2005). Interpreting age effects in second language acquisition. In J. Kroll (Ed.), *Handbook of bilingualism: Psycholinguistics perspectives* (pp. 109-127). Oxford: Oxford University Press.

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced late learners. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 133-159). Mahwah, NJ: Lawerence Erlbaum.

Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997) Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19 (4), 447-465.

Bradlow, A. R., Pisoni, D. B., Arakhane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify English /r/ and /l/: IV. Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America*, 101 (4), 2299-2310.

Chan, A. Y. W. (2014) The perception and production of English speech sounds by Cantonese ESL learners in Hong Kong. *Linguistics*, 52 (1), 35-72.

Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility. *Studies in Second Language Acquisition*, 19 (1), 1-16.

Ferguson, A., & Chowdhury, M. (1960). The phonemes of Bengali. *Linguistic Society of America*, 36 (1), 22-59.

Flege, J. E. (1987). The production of 'new' and 'similar' phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, 15, 47-65.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. *Journal of the Acoustical Society of America*, 89 (1), 70-79.

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

Flege, J. E. (2003). Assessing constraints on second-language segmental production and perception. In A. Meyer and N. Schiller (Eds), *Phonetics and phonology in language comprehension and production, differences and similarities* (pp. 319-355). Berlin: Mouton de Gruyter.

Flege, J. E., & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, 15, 67-83.

Flege, J. E., Munro, M. J., & MacKay, I. R. A. (1995). Factors affecting strength of perceived foreign accent in a second language. *Journal of the Acoustical Society of America*, 97 (5), 3125-3134.

Guion, S. G., Flege, J. E., Akahane-Yamada, R., & Pruitt, J. C. (2000). An investigation of current models of second language speech perception: The case of Japanese adults' perception of English consonants. *Journal of the Acoustical Society of America*, 107 (5), 2711-2724.

Hanulíková, A., Dediu, D., Fang, Z., Bašnaková, J., & Huettig, F. (2012). Individual differences in the acquisition of a complex L2 phonology: A training study. *Language Learning*, 62 (Supplement s2), 79-109.

Higgins, C. (2003). 'Ownership' of English in the outer circle: An alternative to the NS-NNS dichotomy. *TESOL Quarterly*, 37 (4), 615-644.

Jehma, H., & Phoocharoensil, S. (2014). L1 transfer in the production of fricatives and stops by Pattani-Malay learners of English in Thailand. *Asian Social Science*, 10 (7), pp. 67-78.

Jenkins, J. (2005). Implementing an international approach to English pronunciation: The role of teacher attitudes and identity. *TESOL Quarterly*, 39 (3), 535-543.

Jenkins, J. (2006). Current perspectives on teaching English World Englishes and English as a lingua franca. *TESOL Quarterly*, 40 (1), 157-181.

Lakshmanan, U., & Selinker, L. (2001). Analysing interlanguage: how do we know what learners know? *Second Language Research*, 17 (4), 393-420.

Lenneberg, E. H. (1969). On explaining language. *Science*, 164, 635-643.

Long, M. H. (2007). *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum Associates.

Montrul, S. (2010). How similar are adult second language learners and Spanish heritage speakers? Spanish clitics and word order. *Applied Linguistics*, 31 (1), 167-207.

Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, 44 (1), 1-35.

Oyama, S. (1976). A sensitive period for the acquisition of a nonnative phonological system. *Journal of Psycholinguistic Research*, 5 (3), 261-283.

Patowski, M. S. (1990). Age and accent in a second language: A reply to James Emil Flege. *Applied Linguistics*, 11 (1), 73-89.

Rochet, B. L. (1995). Perception and production of second-language speech sounds by adults. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 379-410). Timonium, MD: York Press.

Rohena-Madrazo, M. (2013). Perceptual assimilation of occluded voiced stops by Spanish listeners. In C. Howe et al., (Eds.), *Selected Proceedings of the 15th Hispanic Linguistics Symposium*, (pp. 140-156). Somerville, MA: Cascadilla Press.

Rothman, J. (2008). Why all counter-evidence to the Critical Period Hypothesis is not equal or problematic. *Language and Linguistics Compass*, 2.6, 1063-88.

Scovel, T. (1969). Foreign accents, language acquisition, and cerebral dominance. *Language Learning*, 19 (3-4), 245-253.

Scovel, T. (2000). A critical review of the critical period research. *Annual Review of Applied Linguistics*, 20, 213-223.

Snow, C. E., & Hoefnagel-Höhle, M. (1977). Age differences in the pronunciation of foreign sounds. *Language and Speech*, 20 (4), 357-365.

Strange, W., & Shafer, V. L. (2008). Speech perception in second language learners: The re-education of selective perception. In J. G. Hansen Edwards and M. L. Zampini (Eds.), *Phonology and second language acquisition* (pp. 153-191). Philadelphia: John Benjamins.