Discrimination of English and Thai words ending with voiceless stops by native Thai listeners differing in English experience

Kimiko Tsukada

Department of International Studies
Macquarie University, Australia
Kimiko.Tsukada@humn.mq.edu.au

Rungpat Roengpitya

College of Religious Studies
Mahidol University, Thailand
rungpat@gmail.com

This study examines the discrimination of words ending with voiceless stops /p t k/ in first language (L1) and second language (L2) by three groups of native Thai participants. These participants differed in their countries of residence and experience with L2 English in the formal education system. The first group (T1) was a group of 18 Thai listeners who were living in Australia. The second and third groups consisted of 12 university students (T2) and 12 high-school students (T3) living in Thailand. Eighteen Australian English (AusE) listeners were included as controls. English and Thai words minimally contrasting in the final stop (e.g. 'cap' vs. 'cat') were presented to the listeners to investigate whether L2 phonetic learning occurs even for the contrasts that are expected to be discriminated with high accuracy from the onset of L2 learning and if so, how it is influenced by L2 experience. All three Thai groups showed reasonably accurate discrimination for both English and Thai words, but only T1 showed discrimination accuracy comparable to AusE in English. Further, only T3 was clearly more accurate in discriminating unreleased Thai than English stop contrasts, most of which were accompanied with release bursts. These two findings are taken to be evidence for phonetic learning of specific aspects of L2 contrasts rather than positive L1 transfer.

1 Introduction

Second language (L2) speech research generally focuses on the production/perception of L2 sounds that are in some way unfamiliar to the learner. For instance, English sounds such as /θ/, /ð/ or contrasts like /l/–/r/ unfamiliar to Japanese learners (e.g. Best & Strange 1992; Guion et al. 2000a; Lambacher et al. 2001; Aoyama et al. 2004) or familiar contrasts in unfamiliar contexts (e.g. Broersma 2005) have been examined. This focus is due, possibly, to the assumption that if the new sound systems to be acquired are entirely familiar to the L2 learner, no new learning is required and no difficulty is anticipated (Lado 1957). However, speech is rich in variability and even when there is an apparent mapping between the first language (L1) and L2 sound systems, there may be measurable cross-linguistic phonetic differences. Such differences in voice onset time (VOT) values, for example,

Journal of the International Phonetic Association (2008) 38/3
doi:10.1017/S0025100308003575

may influence learners' L2 speech processing as evidenced in foreign-accented speech (e.g. Flege 1987; Sancier & Fowler 1997; Flege, MacKay & Meador 1999; MacKay et al. 2001).

Piske, Flege & MacKay (2001) reviewed a number of previous studies that examined an overall degree of foreign accent and identified various factors which affect L2 pronunciation accuracy. These include age of L2 learning, length of residence (LOR), amount of L1 use, learners' L1 backgrounds and so forth. Of these factors, LOR has been frequently used as an index of L2 experience. A review of existing literature provides mixed results for the LOR effect on L2 speech learning by adult learners. A number of studies have shown a significant effect of LOR on the perception and/or production of L2 sounds (e.g. Flege, Takagi & Mann 1996; Flege, Bohn & Jang 1997; Guion et al. 2000a; Flege & Liu 2001) and on an overall degree of foreign accent (Flege & Fletcher 1992). That is, in general, more experienced L2 learners with a longer LOR tend to approximate phonetic norms of the target language to a greater extent than do less experienced L2 learners with a shorter LOR. However, there are also studies showing a lack of LOR effect (e.g. Oyama 1976; Flege 1988; Piske et al. 2001; Tsukada et al. 2004, 2005; Flege et al. 2006). Furthermore, in L2 studies with adult migrants, it is often the case that LOR is confounded with participants' age of arrival (AOA) in the host country. This makes it difficult to assess an independent effect of LOR. As described below, the present study attempted to evaluate the effect of L2 learning independently of LOR effects by including groups of participants who were learning English in their home country and did not differ in LOR (and, by extension, AOA).

The aim of the present study was to investigate whether L2 phonetic learning occurs even for contrasts that are expected to be discriminated with high accuracy from the onset of L2 learning. This expectation was motivated by the consideration that, despite cross-language phonetic differences, the contrasts being examined were not unknown to L2 learners due to shared phonetic structures between L1 and L2. We investigate this issue by testing how Thai speakers discriminate English and Thai words minimally contrastive in the place of articulation of the final stops /p t k/. Eighteen Australian English (AusE) speakers were included as controls for the discrimination of word-final stops in English.

1.1 English and Thai stops

This section provides a phonetic description of English and Thai oral stops that is of relevance to the present study. We consider it legitimate to compare the English and Thai final stops with each other to the extent that voiceless stops represented as /p t k/ form a natural class of sounds that occur very frequently in many of the world's languages. According to Abramson & Tingsabadh (1999: 112), 'Thai and English both have word-final voiceless stop consonants at three oral places of articulation that can be phonetically very similar'. They then give an example of the perceptual similarity of the monosyllabic [hɑt̚] for a bilingual speaker who knows English and Thai. Even though the two languages are genetically unrelated, given the phonetic and perceptual similarity observed by experienced phoneticians, it would not be illogical to compare how the two sets of stops might be related in Thai speakers' long-term cognitive representations.

Further support for the legitimacy of this cross-language comparison can be found in the data obtained from Thai speakers learning English (Hancin-Bhatt 2000). In the production task, simple codas consisting of voiceless stops in English nonwords (e.g. 'geet') were accurately reproduced 88% of the time (58 out of 66 tokens). Of the eight instances of inaccurate productions, the most frequent type of error was substitution of stops with fricatives (/t/ → [s] ($n = 2$), /k/ → [f] ($n = 4$), /k/ → [s] ($n = 1$)). Although this pattern cannot be predicted from the native Thai phonological grammar, assuming that the simple codas were naturally released in the stimuli, Thai speakers' substitution errors may be interpreted as the result of their attending to the frication noise in the release burst of the final stop which may be novel and perceptually salient to them. Also, English words such as *cap*, *boot* and *coke* retain their original voiceless stops when borrowed into Thai (Kenstowicz & Suchato 2006), suggesting a correspondence between the two stop systems. Needless to say, it would be desirable in future research to assess empirically their cross-language perceptual assimilation patterns.

Thai oral stops include /p^h t^h k^h p t k ʔ b d/ (Tingsabadh & Abramson 1993) in contrast with /p t k b d g/ for English (e.g. Ladefoged 2006). In the initial and medial positions, Thai has a three-way voicing contrast (voiceless aspirated, voiceless unaspirated, voiced) at bilabial /p^h p b/ and alveolar /t^h t d/ places of articulation (but only aspirated and unaspirated categories /k^h k/ at velar place of articulation), unlike a two-way voicing contrast used in English. Although Thai has no voicing contrast in word-final position (Abramson 1972; Hancin-Bhatt 2000), it permits a glottal stop /ʔ/ in addition to /p t k/ in monosyllabic morphemes with short vowels and, thus, has a four-way place contrast. English, on the other hand, maintains a voicing contrast word-finally. As English and Thai share three voiceless final stops /p t k/, it would be interesting to compare the participants' discrimination of the contrasts involving those native and non-native stops. Since we focused only on the three shared phonemes of the two languages, i.e. /p t k/, it should be pointed out that the contrasts tested are in no way exhaustive for either English or Thai.

Even though the English and Thai final stops may not be phonetically identical to each other, i.e. the stops from the two languages may use slightly different articulators and/or places of articulation and so forth, it is plausible that the three stop contrasts of interest (/p/–/t/, /p/–/k/, /t/–/k/) would be perceived in terms of two distinct phonetic categories by native listeners of both languages (i.e. two-category assimilation according to the Perceptual Assimilation Model, Best 1995). In other words, this study focused on sound contrasts that were phonemically contrastive in both L1 and L2, but that differed substantially in phonetic realization, i.e. sometimes unreleased in English (e.g. Byrd 1993; Lisker 1999; Bent & Bradlow 2003) and always unreleased (or never released audibly) in Thai (e.g. Tingsabadh & Abramson 1993; Abramson & Tingsabadh 1999; Lisker 1999; Kenstowicz & Suchato 2006). Thus, the crucial difference between the word-final stops in English and Thai is that the stops may (but not necessarily) be accompanied by a release burst in the former but must not be in the latter.

Several factors are known to influence the frequency of final release bursts in English stops, including the identity of the preceding vowel (Parker & Walsh 1981; Lisker 1999), the gender of the talker (Byrd 1992, 1993, 1994), place of articulation (Crystal & House 1988; Byrd 1993), dialect (Byrd 1992), speaking style (Bond & Moore 1994; Picheny, Durlach & Braida 1985, 1986) and the position of the stop within the utterance (Halle, Hughes & Radley 1957). The presence of a release burst is not phonemically distinctive in English or in any other language. Despite the absence of a contrastive role for release bursts, unreleased English stops are less intelligible than their released counterparts (e.g. Householder 1956; Malécot 1958; Wang 1959). However, the intelligibility of unreleased English stops would certainly be susceptible to top–down processes in speech perception and/or contextual information.

As for unreleased Thai stops, previous research showed that their intelligibility depended on listeners' language backgrounds and native (but not non-native) listeners were able to identify or discriminate their L1 final stops accurately (Abramson & Tingsabadh 1999; Tsukada 2006). In particular, Abramson and Tingsabadh (1999) interpreted between-group differences in the identification of stops in terms of cross-language differences in the articulatory closing gestures between English and Thai, i.e. there may be more information in the formant transitions for the Thai than for English stops to compensate for the absence of release in the former and the English speakers may not be accustomed to decoding that information as effectively as the Thai listeners.

1.2 Cross-language speech perception

Cross-linguistic comparisons between English and Thai final stops as outlined above, together with the Contrastive Analysis Hypothesis (CAH; Lado 1957), which predicts L2 learning difficulty for areas that differ from L1, might lead to the expectations that both AusE and Thai listeners would discriminate stop contrasts in English and Thai accurately. Since all three contrasts are phonological contrasts in both languages, native phonology should not hinder the discrimination of non-native English stop contrasts by the Thai listeners nor non-native Thai stop contrasts by the AusE listeners.

However, one study (Hallé, Best & Levitt 1999) demonstrated that the L1 phonology could overestimate listeners' responses to non-native sound categories and interfere with the positive transfer from L1 to L2. Four approximants, /w j r l/, occur in French with varying degrees of similarity to the corresponding English sounds. If the equivalence at the traditional phonological level predicts cross-linguistic perception patterns, French listeners would not have difficulties with these English sounds. However, they were found to have some perceptual difficulties with the English /r/. This finding was attributed to the marked articulatory-phonetic differences between the English and French /r/ (i.e. phonetically realized as a central approximant in English and a uvular fricative in French). Of the three contrasts tested (/w/–/j/, /r/–/l/, /w/–/r/), the French listeners had most difficulty with /w/–/r/ and tended to hear the English /r/ as /w/-like.

In previous studies which investigated the effect of L1 on native and non-native stop contrasts (Tsukada 2006; Tsukada & Ishihara 2007), unreleased Thai stop contrasts were discriminated less accurately by AusE and Japanese listeners than by native Thai listeners, implying that different acoustic cues were available to those non-native listeners. The AusE listeners' less than optimal discrimination of Thai stops despite the occurrence of unreleased stops in their L1 seems consistent with the 'phonetic-superiority hypothesis' (Cho & McQueen 2006). According to this hypothesis, unreleased stops are harder to process regardless of the phonological system of the listener's native language, because they contain less acoustic information. The alternative hypothesis, the 'phonological-superiority hypothesis', would predict better performance for non-native sounds that conform to native phonological constraints.

Cho & McQueen (2006) obtained two sets of perception data, i.e. reaction time and phoneme detection accuracy, to test these two hypotheses. They tested Korean and Dutch listeners whose L1s exhibited different phonological constraints with respect to final stops. The listeners heard Dutch and English stimuli with and without release bursts. The Korean listeners responded faster to the unreleased stimuli (phonologically viable, because final stops are unreleased in Korean when produced in isolation, but are phonetically less informative) in listening to both familiar English and unfamiliar Dutch sounds. They were, however, more accurate in detecting the released English (but not Dutch) stimuli even though they were phonologically non-viable (but phonetically more informative). The Dutch listeners, on the other hand, were significantly faster in responding to the released than unreleased stimuli in Dutch, but not in English. Accuracy was affected by the stimulus languages such that the Dutch listeners were more accurate in detecting the unreleased Dutch than English stimuli. There was no language effect for the released tokens and they were highly accurate in both Dutch and English.

1.3 Motivation for the present study

The aim of the present study was to investigate the extent to which L2 phonetic learning may be influenced by L2 experience. To this end, we examined the ability of three groups of native Thai listeners differing primarily in English experience and a control group of 18 AusE speakers to discriminate stop place contrasts (/p/–/t/, /p/–/k/, /t/–/k/) at the end of English and Thai CVC words. Thus, for the Thai groups, the effect of L2 experience on both L1 and L2 speech perception was assessed. Of the three groups, the first (T1) was a group of 18 Thai listeners who were living in Australia; the second and third groups consisted of 12 university students (T2) and 12 high-school students (T3) living in Thailand.

As the Thai participants in this study differed in L2 experience in the formal education system in their home country on the one hand and in LOR on the other hand (with only T1 listeners living in predominantly English-speaking environments in Australia), the design of this study also enabled us to address the theoretical issue of whether L1 perception changes as a result of L2 experience, i.e. plasticity in speech perception. Every Thai participant tested in Australia had an AOA greater than 12. If one's L1 phonological system is firmly established by that age such that no subsequent linguistic experience can alter the existing system, we would not expect to see any difference in L1 perception for the three groups of Thai participants. If, however, the extent of L2 learning experience affects individuals' cognitive representations, we might expect to see T1 showing a better discrimination of L2 English stop contrasts AND poorer discrimination of L1 Thai stop contrasts than T2 and T3, native Thai listeners living in Bangkok. Although all participants have studied English at school, their daily use of English is expected to be much greater for T1 than for T2 and T3 participants.

If we apply the characterization in Cho & McQueen (2006) reviewed above to the present study, the released English stops are phonologically non-viable but phonetically enriched for the Thai listeners, whereas the unreleased Thai stops are NOT non-viable but phonetically less informative for the AusE listeners. However, there was a clear asymmetry in language familiarity in that Thai was unknown to the AusE listeners, but English was not to the Thai listeners. It is thus possible that the English released stops were no longer a serious violation for some of the Thai listeners depending on their experience with English.

2 Method

The purpose of this experiment was to assess the discrimination of English and Thai words ending with voiceless stops by three groups of Thai listeners differing in English language experience and a group of AusE listeners.

2.1 Speech materials

The English and Thai test words are given in appendices A and B. Words ending with stops minimally contrastive in places of articulation (e.g. 'sip' vs. 'sit') were used as stimuli. A total of 117 and 125 unique waveforms for English and Thai, respectively, were arranged in triads and presented via headphones at a self-selected comfortable level using a notebook computer. The same stimuli were used in previous studies (Tsukada 2006; Tsukada & Ishihara 2007). Note that the context vowels were different for the English and Thai stimuli. For the Thai stimuli, short vowels /i a u/ were used while a much wider range of vowels, both short and long, were used for English.[1]

English and Thai monosyllabic CVC words were read by native speakers of Australian English and Thai in the MARCS Auditory Laboratories recording studio at the University of Western Sydney, Australia. Test words (all real words printed in English or Thai) were presented visually to each speaker in randomized orders on the computer screen, one word at a time. All Thai words were written using conventional Thai script, and they had either high or low tones. Words with two different tones were collected so that the discrimination of tokens differing solely in lexical tones could be examined in a separate study. In order to

[1] Vowel quality and quantity need to be more closely matched for the two languages in future research, as there may be cross-linguistic differences in how vowel identity interacts with coarticulatory effects. Specifically, we acknowledge that there may be differences in the coarticulatory effects between English short and long vowels before different stops.

prevent the speaker gender from affecting listeners' responses, male and female voices were not presented together within a given trial.[2] Instead, tokens from only one gender (tokens produced by three female speakers for the English stimuli and tokens produced by three male speakers for the Thai stimuli) were used.[3] In each trial, the recordings of these three speakers were used throughout in different orders for each language.

The recorded speech materials were digitized at 44.1 kHz using the CoolEdit program and the amplitude of each sound file was normalized to 50% of the peak following the procedures used in previous research (e.g. Guion et al. 2000a; Aoyama et al. 2004; Flege & MacKay 2004). All test words were segmented and stored in separate files. More than 90% (107 out of 117 unique waveforms) of English final stops were produced with an audible release burst although speakers were not given specific instructions as to how the final stops should be pronounced. The ten unreleased tokens included three instances of /p/, six instances of /t/ and one instance of /k/.

2.2 Listeners

Three groups of native Thai listeners participated. The first group (T1) included 18 native Thai speakers (4 male, 14 female) with a mean age of 33.1 years (*sd* = 11.4, *range* = 20–57) who lived either in Sydney or in Canberra, Australia, at the time of testing. They responded to an advertisement placed in a local newspaper. The T1 group's mean LOR in Australia was 7.1 years (*sd* = 6.4, *range* = 2.5–30.3). Their mean AOA in Australia was 25.4 years (*sd* = 7.6, *range* = 14–41). These listeners were tested in the MARCS Auditory Laboratories, University of Western Sydney or at the Australian National University in Canberra by the first author. The experimental sessions with the T1 group were conducted in English.

The second (T2) and third (T3) groups consisted of 12 (2 male, 10 female) undergraduate university students at Chulalongkorn University and 12 female high-school students, respectively, living in Thailand. The T2 and T3 groups were recruited by the second author and her assistant from the student population on the basis of their availability at the time of testing. These participants were not required to satisfy any special conditions. The listeners in the T2 and T3 groups were tested in the Department of Linguistics at Chulalongkorn University by the second author. The experimental sessions with the T2 and T3 groups were conducted in Thai.

These two groups did not differ in LOR, as the participants in both groups had never lived in English-speaking countries (i.e. LOR = 0).[4] The two groups did, however, differ in their chronological age. T2 had a mean age of 19.3 years (*sd* = 0.9, *range* = 18–20) and T3, 17.6 years (*sd* = 0.9, *range* = 15–18). Although both groups indicated that they had studied English for nearly the same number of years (T2: 12.1 vs. T3: 12.7 years),[5] only the T2 listeners were receiving English education at university.

In a non-English-speaking country such as Thailand, students may start learning English as their L2 at primary school and/or because their parents want them to. For instance, some students in this study began their first English lesson as early as 3 years of age, i.e. prior

[2] For example, if a trial such as 'sip$_{m1}$'–'sip$_{f2}$'–'sick$_{f3}$' were presented (where the subscripts indicate different talkers, m being male and f, female), a listener may 'wrongly' select the first token rather than the third token to be an odd item.

[3] Although we do not expect cross-gender phonetic effects to differ substantially across listener groups, it would have been desirable to include both male and female talkers for both stimulus languages. This inconsistency was unavoidable at the time of stimuli preparation.

[4] The only exception was a student in the T2 group who spent one academic year in the US as an exchange student. We decided to retain this subject who was accidentally included in the study as her response patterns did not substantially differ from those of the rest of the participants in T2.

[5] This suggests that the high-school students (T3) started learning English prior to their formal schooling, which resulted in more years of English learning experience despite their younger age.

to formal schooling. However, it is difficult to assess the efficacy of unsystematic English education. Thus, we only considered the number of years the participants in the T2 and T3 groups spent learning English within the formal education system as their English language experience. By comparing these two groups, the effect of English experience could be assessed independently of the LOR effect.

In an attempt to gain an insight into the nature of their English input, we asked four participants each in the T2 and T3 groups (i) how many hours of English classes they had per week, (ii) how many teachers they had for English classes, (iii) how many of those teachers were native speakers of Thai, (iv) how many of those teachers were native speakers of English, and (v) which country the native English-speaking teachers were from. On average, the T2 group had 8 hours of English classes per week and two or three teachers; one or two of these teachers were native speakers of Thai and one was a native speaker of English. Native English-speaking teachers were from the US and the UK. The T3 group, on the other hand, had 6.8 hours of English classes per week and three or four teachers; two or three of these teachers were native speakers of Thai and one was a native speaker of English. Native English-speaking teachers were from the US and the UK. From this informal survey, it can be inferred that the T2 and T3 participants had a mixture of native and Thai-accented English input.

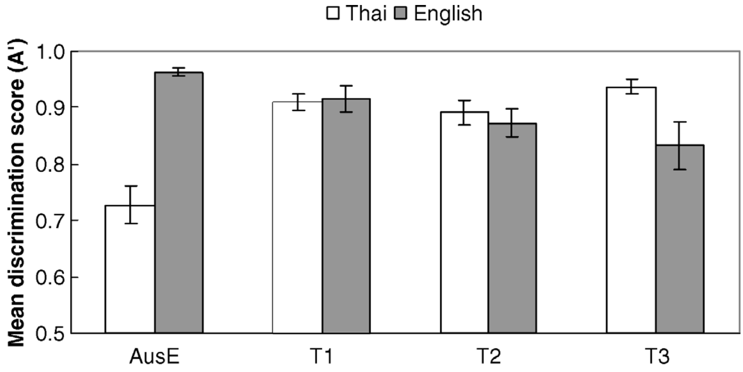
In addition to the three Thai groups, a group of 18 AusE listeners (7 male, 11 female) were recruited as controls. The AusE participants were students at the University of Western Sydney or Macquarie University. Three AusE listeners tested in the MARCS Auditory Laboratories at the University of Western Sydney were enrolled in a first-year Psychology course, and 15 AusE listeners tested in the Department of Linguistics at Macquarie University were enrolled in a first- or second-year Linguistics course. They received credit points for the respective courses for their participation. None of them had any knowledge of Thai. They had a mean age of 23.5 years (*sd* = 8.3, *range* = 18–48). Each participant was tested individually in a session lasting between 45 and 60 minutes. According to self-report, all participants had normal hearing and did not have any language deficiency in their L1s.

2.3 Procedure

A categorial discrimination test (CDT) employed in previous L2 speech research (e.g. Flege et al. 1999; Guion et al. 2000a; Flege 2003; Wayland & Guion 2003, 2004; Aoyama et al. 2004; Flege & MacKay 2004; Tsukada et al. 2005) was used in this study. Monosyllabic CVC words ending with stop tokens were presented in triads to each listener whose task was to identify an odd item if any. As described in Wayland & Guion (2003: 118), this is 'a version of ABX discrimination task' and 'is designed to minimize response bias (guessing)'. We considered this test to be suitable for examining phonetic processes used in cross-language speech perception. A high level of performance in this task would require not only the use of purely auditory information but also the establishment of phonetic categories for one or both stops in a given contrast.

Each contrast was tested by change and no-change trials. The three word tokens in all trials were spoken by three different talkers, and so were always physically different even in no-change trials. The listeners were asked to choose a word that was different from the other two, if there was any. The change trials contained an odd item. For example, a change trial testing the /p/–/t/ contrast might consist of 'sip_2'–'sit_1'–'sip_3' (where the subscripts indicate different talkers). The correct response for change trials was the button ('1', '2', or '3') indicating the position of the odd item, which occurred with near-equal frequency in all three possible serial positions. The change trials tested the participants' ability to respond appropriately to relevant phonetic differences between tokens and distinguish stops drawn from two different categories.

The correct response to no-change trials, which contained three different instances of a single category (e.g. $/\text{p}/_1$ $/\text{p}/_3$ $/\text{p}/_2$ or $/\text{t}/_3$ $/\text{t}/_1$ $/\text{t}/_2$), was a fourth button marked 'NO'. The



AusE = Australian English listeners; T1 = Native Thai listeners living in Sydney or Canberra, Australia; T2 = Native Thai university students living in Bangkok, Thailand; T3 = Native Thai high school students living in Bangkok, Thailand.

Figure 1 Mean discrimination scores (A′) for Thai and English stimuli averaged across place contrasts. The brackets enclose ±1 standard error.

no-change trials tested the participants' ability to ignore audible but phonetically irrelevant within-category variation (in e.g. voice quality). The participants were required to respond to each trial, and were told to guess if uncertain. A trial could be replayed as many times as the listener wished, but responses could not be changed once given. The inter-stimulus interval in all trials was 0.5 s. Two blocks of 58 trials were presented with the second block being a repeat of the first block. A different randomization was used for each block. The first ten trials were for practice and were not analyzed. The 48 trials in each block consisted of 36 change trials testing three contrasts (12 trials each for /p/–/t/, /p/–/k/, /t/–/k/) and 12 no-change trials (4 trials each for /p/–/p/, /t/–/t/, /k/–/k/). The English and Thai stimuli were presented in separate blocks.

2.4 Analysis

Responses to the change and no-change trials were used to calculate A-prime (A′) scores (Snodgrass, Levy-Berger & Haydon 1985), an index of discrimination accuracy. These scores were based on the proportion of 'hits' obtained for each contrast and the proportion of 'false alarms'. If the proportion of hits (Hs) equaled the proportion of false alarms (FAs), then A′ was set to 0.5. If H exceeded FA, then $A' = 0.5+((H–FA)*(1+H–FA))/((4*H)*(1–FA))$. However, if FA exceeded H, then $A' = 0.5–((FA–H)*(1+FA–H))/((4*FA)*(1–H))$. An A′ score of 1 indicated perfect sensitivity, whereas an A′ score of 0.5 or lower indicated a lack of sensitivity.

3 Results and discussion

Figure 1 shows discrimination scores for the Thai and English stimuli averaged across contrast types. All three groups of Thai listeners obtained mean discrimination scores higher than 0.80 for both English and Thai stop contrasts. However, the pattern of results varied according to the level of English experience. The T1 and T2 listeners' mean A′ values for English and Thai did not differ significantly: 0.92 vs. 0.91 for T1, 0.87 vs. 0.89 for T2, suggesting that they were competent in discriminating the final stops in both their L1 and their L2. The T3 listeners, on the other hand, were more accurate in discriminating L1 Thai stops than L2 English stops (0.94 vs. 0.83). This suggests that although all the Thai listeners were able to discriminate the English words minimally contrasting in the final stops relatively well, their accuracy depended on their experience with English. The more experienced they were with English, the better adjusted they were with novel phonetic realizations of the final stops. The AusE listeners were much better at discriminating the final stops in English than in Thai (0.96 vs. 0.73), possibly because they were less well attuned than the Thai listeners to the absence of release bursts in the Thai stimuli where there may be 'some kind of articulatory compensation that provides useful auditory enhancement of the closing gesture' (Abramson & Tingsabadh 1999: 113).

A three-way ANOVA with Group (AusE, T1, T2, T3) as a between-subjects factor, and Stimulus language (English, Thai) and Contrast (/p/–/t/, /p/–/k/, /t/–/k/) as within-subjects factors was carried out. The dependent variable was the discrimination score (A′) obtained by each listener. Significant interactions were explored by Tukey's tests.

All three main effects reached significance [Group: $F(3, 56) = 3.4$, $p = .023$, Stimulus language: $F(1, 56) = 5.2$, $p = .0267$, Contrast: $F(2, 112) = 15.3$, $p = .0001$] and so did the Group × Stimulus language, Contrast × Stimulus language and Group × Contrast interactions [G × S: $F(3, 56) = 33.3$, $p = .0001$, C × S: $F(2, 112) = 32.6$, $p = .0001$, G × C: $F(6, 112) = 3.7$, $p = .0021$]. A three-way interaction was also significant [$F(6, 112) = 8.1$, $p = .0001$], apparently because of different patterns of two-way interactions for the two stimulus languages as shown in figure 2(a), (b). Group × Contrast ANOVAs were therefore conducted for each language separately.

3.1 Thai stimuli

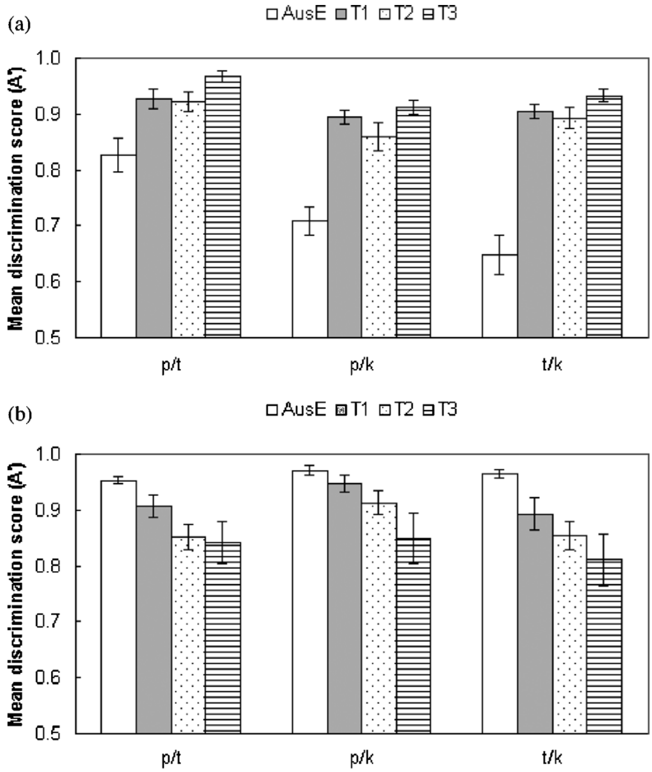
Figure 2(a) shows discrimination scores for the Thai stimuli by the 4 groups of listeners as a function of contrast types. The overall mean discrimination scores were highest for the T3 group (0.94) and lowest for the AusE group (0.73) and intermediate for the T1 (0.91) and T2 (0.89) groups.

Both main effects reached significance [Group: $F(3, 56) = 27.1$, $p = .0001$, Contrast: $F(2, 112) = 27.6$, $p = .0001$] as well as the two-way interaction [$F(6, 112) = 7.7$, $p = .0001$], which was explored through simple effects tests. The simple effect of Contrast was significant for all groups. Although all groups discriminated /p/–/t/ most accurately, the pattern of the Contrast effect was slightly different as is seen in figure 2(a). All three Thai groups showed the same pattern of the Contrast effect, i.e. their discrimination score was highest for /p/–/t/, lowest for /p/–/k/ and intermediate for /t/–/k/. Only the difference between the /p/–/t/ and /p/–/k/ contrasts reached significance for the three Thai groups. For the AusE group, discrimination accuracy was higher for /p/–/t/ than for /p/–/k/ and /t/–/k/, which did not differ from each other. The simple effect of Group was significant for all three contrasts [$F(3, 56) = 7.6$–32.1, $p < .001$]. The advantage of the Thai groups over the AusE group can be clearly seen in figure 2(a) for all stop contrasts tested. There was no statistically significant difference in discrimination accuracy among the three Thai groups for any of the Thai contrasts.

3.2 English stimuli

Figure 2(b) shows discrimination scores for the English stimuli by the 4 groups of listeners as a function of contrast types. The overall mean discrimination scores were highest for the AusE group (0.96), lowest for the T3 group (0.83) and intermediate for the T1 (0.92) and T2 (0.87) groups.

Both main effects reached significance [Group: $F(3, 56) = 6.1$, $p = .0012$, Contrast: $F(2, 112) = 15.2$, $p = .0001$] as well as the two-way interaction [$F(6, 112) = 2.2$,



AusE = Australian English listeners; T1 = Native Thai listeners living in Sydney or Canberra, Australia; T2 = Native Thai university students living in Bangkok, Thailand; T3 = Native Thai high school students living in Bangkok, Thailand.

Figure 2 Mean discrimination scores (A′) showing Group × Contrast interaction. (a) Thai stimuli. (b) English stimuli. The brackets enclose ±1 standard error.

$p = .0471$]. The interaction, which arose possibly due to differing effects of Contrast for each group, was explored through simple effects tests. The simple effect of Contrast reached significance only for the T1 and T2 groups. Although all groups discriminated /p/–/k/ most accurately, the pattern of the Contrast effect was slightly different. For the T1 group, only the difference between /p/–/k/ (most accurate) and /t/–/k/ (least accurate) contrasts reached significance. For the T2 group, discrimination accuracy was higher for /p/–/k/ than for /t/–/k/ and /p/–/t/, which did not differ from each other. The simple effect of Group was significant for all three contrasts [$F(3, 56) = 5.1$–5.9, $p < .01$]. As summarized in table 1, while the T3 group significantly differed from the AusE group for all three contrasts, the T1 group did not differ from the AusE group for any of the contrasts. The T2 group differed from the AusE group for two of the three contrasts tested, showing the intermediate level of performance.

If T1 showed high discrimination accuracy for the English stimuli as a result of positive L1 transfer, T2 and T3 should enjoy the same benefit, as they share the same L1 with T1. These results are more consistent with the view that greater L2 experience led to more successful phonetic learning.

Table 1 Summary of one-way ANOVAs testing for the simple effect of Group for the English stimuli (A′).

Contrast	$F(3, 56)$	p	Tukey's test
/p/-/t/	5.9	.0014	AusE > T2, T3[a]
/p/-/k/	5.1	.0034	AusE, T1 > T3
/t/-/k/	5.7	.0018	AusE > T2, T3[a]

AusE = Australian English listeners; T1 = Native Thai listeners living in Sydney or Canberra, Australia; T2 = Native Thai university students living in Bangkok, Thailand; T3 = Native Thai high school students living in Bangkok, Thailand
[a]Although T1 had higher scores than T2 and T3 for /p/-/t/ and /t/-/k/, the difference did not reach statistical significance. In other words, for these two contrasts, T1 did not significantly differ from any of the three groups.

3.3 Comparison of coarticulatory effects in English and Thai

As mentioned earlier, it would have been desirable to phonetically match the English and Thai stimuli used in the experiment as closely as possible and just allow the final stops to vary for the presence or absence of a release burst. However, due to unfortunate oversight, there were other variations in the two sets of stimuli including the context vowels. In these circumstances, it would not be informative to compare characteristics of the actual stimuli used. Instead, we provide a phonetic description of more comparable vowels – /iː ɐː ʉː/[6] (/i a u/ in the traditional transcription system proposed by Mitchell (1946)) in AusE and /iː aː uː/ in Thai – preceding word-final /p t k/ to gain an insight into coarticulatory effects in the two languages. These materials are different from what was presented to the listeners and are intended for illustrative purposes rather than a comprehensive description of the vowels in the two languages under investigation.

It should be noted here that AusE is a non-rhotic variety. Some phoneticians consider vowel length to be phonemic in AusE (Cochrane 1970). According to Cox (2006b: 149), '[t]he major difference between the long and short vowels is simply one of total duration, however, the difference is relative rather than absolute as contextual and prosodic factors affect the ultimate length of the vowel'.

Figures 3 and 4 present the F1 and F2 trajectories of the AusE /iː ɐː ʉː/ and the Thai /iː aː uː/ produced by two male speakers each to observe coarticulatory effects of the following stops /p t k/. F1 and F2 of the target vowels were automatically tracked using the signal processing package ESPS/Waves (http://www.speech.kth.se/software/). The settings were 12th order linear predictive coding analysis, cosine window, 49-ms frame size, and 5-ms frame shift. The EMU speech database system (http://emu.sourceforge.net/) was used for phonetically labeling the speech segments of interest and the formant values were calculated in the R statistical environment (http://cran.r-project.org/). The beginning and end of each vowel token was identified by inspection of wide-band spectrograms and time domain waveforms. The formant values were extracted at ten equidistant time points throughout the vocalic portion, linearly interpolated and plotted as a function of time normalized to average vowel durations as appropriate.

Following Abramson & Ren (1990), we only included the Thai vowels on the low tone for this presentation. While long Thai vowels /iː aː uː/ may be phonetically more comparable to their English counterparts, this resulted in mostly nonwords. We thus present, in addition, short Thai vowels /i a u/ which contrast phonemically with long /iː aː uː/ (e.g. Gandour 1984; Tingsabadh & Abramson 1993) produced in real words by the same speakers (figure 5). Both real words and nonwords were transcribed using Thai script and for native speakers of the language there was no ambiguity as to how each word should be pronounced. Only real words

[6] The AusE vowels /iː ɐː ʉː/ (or /i a u/) represent vowels in *heed*, *hard* and *who'd*, respectively. The use of these revised phonetic symbols was proposed by Harrington, Cox & Evans (1997) and adopted in more recent work (Cox 2006a, b).

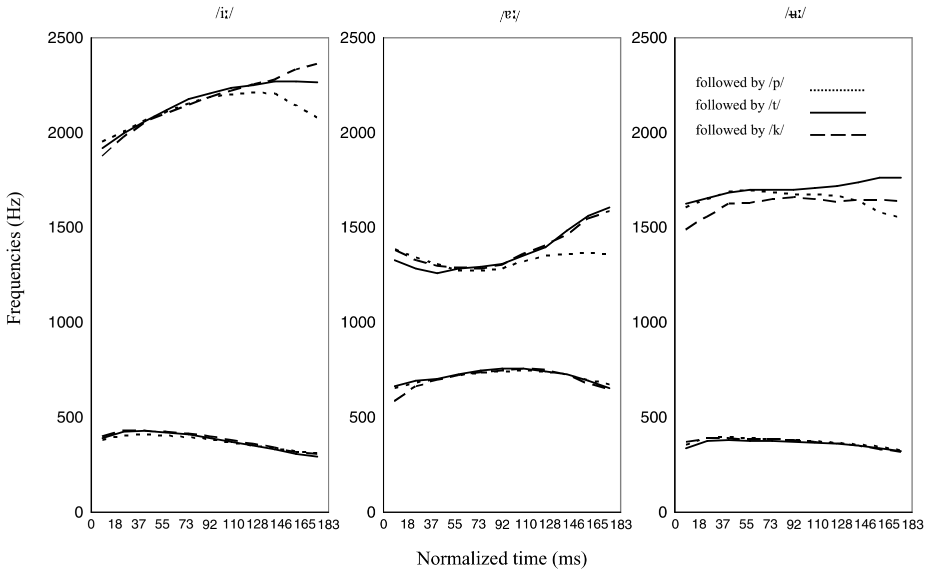


Figure 3 Formant trajectories of Australian English vowels /iː ɐː ʉː/ preceding /p t k/ spoken by two male speakers.

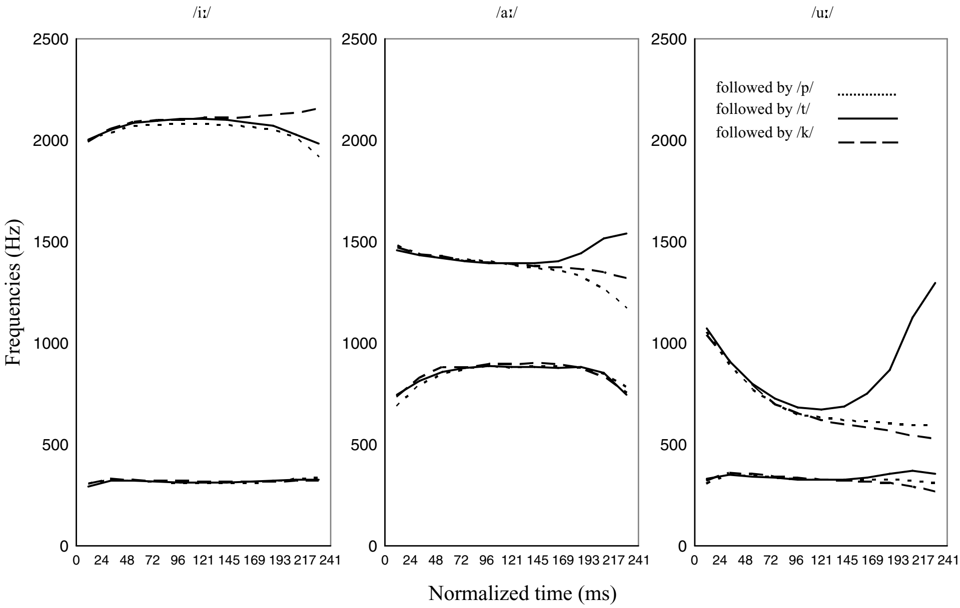


Figure 4 Formant trajectories of long Thai vowels /iː aː uː/ preceding /p t k/ spoken by two male speakers.

were used for English. The initial consonants included /p k f h s ʃ tʃ dʒ l m w/ for English and /pʰ tʰ kʰ p t k b d f h cʰ c/ for Thai.

It has to be pointed out that the selection of the three representative vowels, i.e. /iː ɐː ʉː/ for AusE and /iː aː uː/ for Thai, was one of convenience and we would expect to observe cross-language phonetic differences in these two sets of vowels. For instance, it is clear in

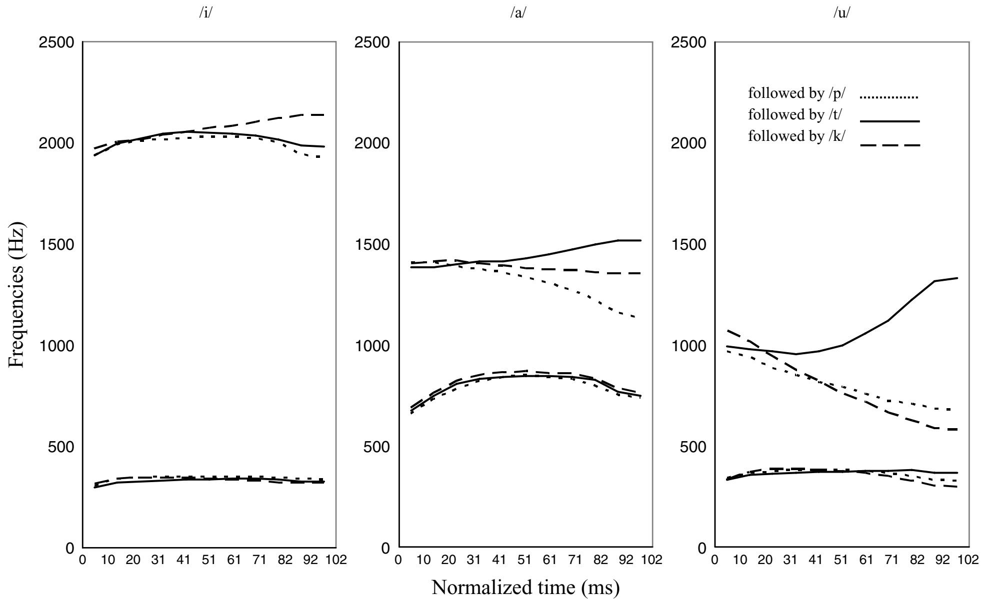


Figure 5 Formant trajectories of short Thai vowels /i a u/ preceding /p t k/ spoken by two male speakers.

figures 3 and 4 that the AusE /ʉː/ and the Thai /uː/ show distinct acoustic characteristics with the former having a much higher F2 than the latter throughout the entire vowel duration. This lends support to the use of the symbol /ʉː/ (for the fronted /uː/) proposed in previous research on AusE vowels (Harrington et al. 1997) and adopted in more recent research (Cox 2006a, b). The raised F2 for /ʉː/ may be due to less lip-rounding and/or more tongue-fronting (Watson & Harrington 1999). The F2 trajectories for the Thai /u–uː/ before /t/ show a sharp upward movement towards the alveolar locus while the fronted AusE /ʉː/ does not show such movement.

The English /iː/ and the Thai /iː/ also showed considerable acoustic differences, with the former having an extensive onglide and being more diphthongal than the latter. This is a well-known phonetic feature of the /iː/ vowel in AusE (Cox 2006b). Cross-language differences appeared more limited for the low vowels except that differentiation of the F2 transitions into the final /t/ and /k/ seen in the Thai /aː/ was absent in the AusE /ɐː/.

In general, it seemed that there was a greater influence of the following stops on the Thai than AusE vowels. This may be because, in the absence of a release burst, the closing gestures of the Thai vowels need to contain sufficient information to signal the place of articulation of the following stops (Abramson & Tingsabadh 1999; Lisker 1999; Kenstowicz & Suchato 2006). A similar conclusion was reached for the unreleased final stops in Cantonese (Ciocca, Wong & So 1994).

It is clear in figures 3 and 4 that the F2 trajectories are considerably influenced by the final stops. However, different stops affected the three vowels differently. The F1 and F2 values at the vowel offset were submitted to a (2) Language × (3) Vowel × (3) Stop ANOVA to assess the extent of coarticulatory effects of the following stops. Separate analyses were conducted for the F1 and F2 values. What is of interest in this analysis is the pattern of interactions involving the Stop factor.

Table 2 summarizes the results of three-way ANOVAs for F1 and F2. For F1, the main effects of Language and Vowel reached significance, but not the effect of Stop. Of the three

Table 2 Summary of three-way ANOVAs for F1 and F2. Significant results are in boldface.

	Factor	*df*	*F*	*p*
F1	**Language (L)**	1, 29	125.2	.0001
	Vowel (V)	2, 58	1016.2	.0001
	Stop (S)	2, 58	2.7	ns
	L × V	2, 58	56.9	.0001
	L × S	2, 58	1.8	ns
	V × S	4, 116	0.6	ns
	L × V × S	4, 116	5.3	.0001
F2	**L**	1, 29	457.4	.0001
	V	2, 58	773.7	.0001
	S	2, 58	194.8	.0001
	L × V	2, 58	80.6	.0001
	L × S	2, 58	41.9	.0001
	V × S	4, 116	39.1	.0001
	L × V × S	4, 116	34.8	.0001

Table 3 Summary of two-way ANOVAs for each language for F1 and F2. Significant results are in boldface.

	Factor	*df*	*F*	*p*
English F1	**Vowel (V)**	2, 18	1129.6	.0001
	Stop (S)	2, 18	0.4	ns
	V × S	4, 36	5.7	.0011
English F2	**V**	2, 18	146.5	.0001
	S	2, 18	50.1	.0001
	V × S	4, 36	0.8	ns
Thai F1	**V**	2, 40	880.9	.0001
	S	2, 40	5.3	.0094
	V × S	4, 80	3.6	.0066
Thai F2	**V**	2, 40	1054.5	.0001
	S	2, 40	279.5	.0001
	V × S	4, 80	150.6	.0001

two-way interactions, only the Language × Vowel interaction reached significance. This is highly expected, as there would be cross-language differences between the English and Thai vowels in their phonetic realizations. A three-way interaction also reached significance. For both languages, F2 showed a much greater effect of the final stops as is clearly seen in figures 3 and 4. All main and interaction effects reached significance. A significant three-way interaction arose presumably due to different patterns of the Vowel × Stop interaction for Thai and AusE. Therefore, two-way (Vowel × Stop) ANOVAs were conducted separately for each language (table 3). Significant interactions were explored by the simple effects tests and the results are presented in table 4.

For AusE, only the main effect of Vowel and the two-way interaction reached significance for F1. The simple effect of Vowel was significant for all stops as expected. The simple effect of Stop was significant for /iː/ and /ɐː/. For /iː/, the F1 offset was significantly higher

Table 4 Summary of the simple effect of Stop for each language for F1 and F2.

	Vowel	*df*	*F*	*p*	Tukey's test
English F1	/iː/	2, 158	7.0	.0013	p, k > t
	/ɐː/	2, 56	9.6	.0003	p > t, k
Thai F1	/iː/	2, 58	5.4	.0071	p > k
	/uː/	2, 40	28.3	.0001	t > p > k
Thai F2	/iː/	2, 58	43.7	.0001	k > t, p
	/aː/	2, 50	141.8	.0001	t > k > p
	/uː/	2, 40	465.9	.0001	t > p, k

before /p/ (310.1 Hz) and /k/ (308.3 Hz) than before /t/ (293.6 Hz). For /ɐː/, the F1 offset was significantly higher before /p/ (678.6 Hz) than before /t/ (604.6 Hz) and /k/ (591.0 Hz). As for F2, both main effects reached significance, but the two-way interaction did not. This suggests that the effect of the following stops was comparable for all three vowels. Averaged across three vowels, the F2 offset was significantly higher before /k/ (1854.6 Hz) and /t/ (1822.8 Hz) than before /p/ (1501.6 Hz).

For Thai, both main effects and the two-way interaction reached significance for F1 and F2. The simple effect of Vowel was significant for all stops as expected, but the simple effect of Stop was significant only for /iː/ and /uː/ for F1. For /iː/, the F1 offset was significantly higher preceding /p/ (336.5 Hz) than preceding /k/ (319.7 Hz). For /uː/, the F1 offset was significantly higher before /t/ (357.2 Hz) than before /p/ (317.0 Hz) and, in turn, it was higher before /p/ than before /k/ (274.7 Hz). As for F2, the simple effect of Stop was significant for all vowels. For /iː/, the F2 offset was significantly higher before /k/ (2150.4 Hz) than before /t/ (1994.9 Hz) and /p/ (1952.6 Hz). For /aː/, the F2 offset was significantly higher before /t/ (1542.3 Hz) than before /k/ (1339.4 Hz) and, in turn, it was higher before /k/ than before /p/ (1206.2 Hz). Finally, for /uː/, the F2 offset was significantly and substantially higher before /t/ (1242.3 Hz) than before /p/ (590.3 Hz) and /k/ (529.5 Hz). While many of the differences in the formant values at the vowel offset reached statistical significance, some of them may be of little perceptual relevance. For instance, in both languages, the difference between the highest and lowest F1 values for /iː/ was less than 20 Hz.

A comparison of figures 4 and 5 shows that, in relation to vowel length, the consonantal influence on the vocalic portion was greater in short than long Thai vowels. Long Thai vowels were, on average, more than twice as long as their short counterparts (241 ms vs. 102 ms). While the F2 trajectories started to diverge prior to the temporal midpoint for the short vowels, the diversion did not start until later for the long vowels. In other words, F2 of the short vowels showed constant, dynamic movement for most of the vocalic portion, but such movement was localized to the second half of the vocalic portion for the long vowels. This may be because the articulators need to start moving towards the final stops earlier in short vowels in order to achieve the closure.

It has been reported that the vowel length contrast in Thai is primarily cued by vowel duration with some vowel quality differences between the two length categories: the short vowels tend to be more open in the vowel space than the long vowels (Abramson & Ren 1990). Figure 6 shows ellipse plots of the averaged F1 and F2 values extracted at the vowel midpoint. Each ellipse contains approximately 95% of the data points. A direct comparison of the short and long Thai vowels lends partial support to the observation made by Abramson & Ren (1990) in that the short /u/ occupied a lower and more central portion of the vowel

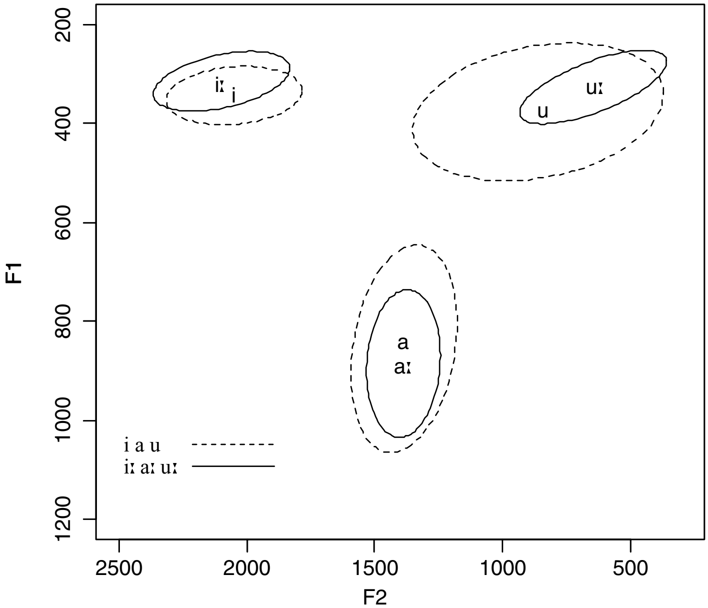


Figure 6 Ellipse plots of short (/i a u/) and long (/iː aː uː/) Thai vowels (in Hz) at the vowel midpoint spoken by two male speakers. Each ellipse contains approximately 95% of the data points.

space than the long /uː/. For the /a–aː/ and /i–iː/ pairs, vowel length did not appear to have a substantial influence on vowel quality at least for this set of data.

4 General discussion

The present study examined the discrimination of English and Thai words ending with voiceless stops by three groups of native Thai participants differing in their English experience and a group of native speakers of Australian English. As mentioned in the Introduction regarding the familiarity with the non-native stimuli, it needs to be pointed out that the experimental tasks were not equivalent for the Thai and AusE groups. This lack of experimental control was due to the fact that the Thai lexicon mostly consists of monosyllabic words, which motivated us to use real words as stimuli. For the AusE listeners responding to the Thai stimuli, the task was almost purely phonetic, as the Thai words were virtually nonwords to them. For the Thai listeners responding to the English stimuli, on the other hand, the task involved phonetic processing as well as word discrimination in their L2. However, we argue that, by presenting the SAME set of stimuli under controlled conditions to all listeners who differ in their L1 and experience with L2, variation in their perceptual performance is still interpretable, as it ought to be attributed to those variables that define the profile of the listeners.

There were three main findings. Firstly, although all three Thai groups were able to discriminate both English and Thai stop contrasts reasonably well, it was only the T1 listeners who discriminated the English final stop contrasts as accurately as the AusE listeners. While the T1 and T2 groups showed balanced discrimination accuracy in the two languages (T1: 0.92 for English and 0.91 for Thai, T2: 0.87 for English and 0.89 for Thai), T3 showed better discrimination accuracy for the Thai (0.94) than English contrasts (0.83). This suggests that the T1 (and T2) listeners, with more years of English experience, learned subtle phonetic differences between the L1 and L2 sounds to a greater extent than did the T3 group, who had least English experience.

In order to gain support for the interpretation that the between-group differences in our results were more likely resulting from the differences in listeners' perceptual abilities than from the differences in other factors (such as their familiarity with the test words), we asked four listeners each in the T2 and T3 groups to rate each of the 54 English stimulus words – given below in appendix A (excluding two proper nouns, 'Pete' and 'Coke') – on a scale ranging from 1 ('most familiar') to 5 ('least familiar') after their discrimination abilities were tested. The average score across these eight listeners was 1.8 (1.88 for T2 and 1.73 for T3, respectively). Only 7 out of 54 words were given a score higher than 4. None of the individual mean scores exceeded 2. Thus, we assume that the English words used were reasonably familiar to the Thai listeners including those who were less experienced with English.[7]

The second finding was that there was no significant difference among the three Thai groups in discriminating their L1 stop contrasts. This is despite the fact that the T1 group was living outside their L1 environment for an average of seven years at the time of testing, suggesting that L2 learning did not negatively affect the L1 perception. This finding is contrary to the results of some studies showing the effect of L2 learning on L1 (e.g. Flege 1987; Guion, Flege & Loftin 2000b; Guion 2003).

Our third finding was that the AusE listeners showed highly accurate discrimination for the English stops, but less so for the Thai stops despite the fact that English and Thai share the same phonemic contrasts in stop sounds and the positive transfer from their L1 English may be predicted. This pattern of results may be adequately explained by the 'phonetic-superiority hypothesis' (Cho & McQueen 2006) mentioned earlier. As reviewed in the Introduction, for the AusE listeners, both English and Thai stimuli were phonologically viable. Therefore, if the 'phonological-superiority hypothesis' holds, the AusE listeners should show accurate discrimination for both English and Thai stimuli. Under the 'phonetic-superiority hypothesis', unreleased Thai stops are difficult to process perceptually, as there is less acoustic information in the speech signal. Further, we agree with Cho & McQueen (2006) that language familiarity influences the extent to which listeners exploit acoustic-phonetic cues in non-native perception.

It is difficult to quantify the communicative value of different acoustic information for individuals with different linguistic experience. However, we observed that the AusE and Thai vowels apparently differed in the extent of coarticulatory effects as encoded in the formant transitions. Presumably, in the absence of a release burst in the final stop, Thai speakers need to pay attention to acoustic information contained in the formant transitions more than AusE speakers in the acquisition of their respective L1.

Despite the expectation that the Thai listeners would discriminate English stops accurately due to (i) the shared stop categories with non-identical, but similar places of articulation for their L1 and L2 final stops and (ii) their experience with unreleased L1 stops, it appeared that the T3 listeners have yet to learn to decode an apparently enriched acoustic cue, i.e. clear release burst in the English stops. It is possible that the release burst in English stops provided

[7] Of course, this does not guarantee that their discrimination accuracy would be high, because there would be a difference in cross-language phonotactic constraints. An anonymous reviewer pointed out that the English words *bite* and *bike* might be problematic to the Thai listeners, since comparable short diphthongs in Thai are never checked with any consonant.

extra, but irrelevant, information to the native Thai listeners who were inexperienced with English. It may be that, to the less experienced listeners, the clear burst was even a distractor, as it violates the native phonological constraints (Cho & McQueen 2006) and hindered their accurate discrimination. The present results could be interpreted as suggesting that non-native listeners can overcome a violation of native phonological constraints depending on their linguistic experience. Perhaps, the T3 listeners need more hours or lessons, with a focus on L2 speaking and listening, in addition to their L2 grammar hours in their high-school education before they modify their L1 perceptual behaviors and learn to perceive English final stops efficiently.

To confirm or disconfirm the above hypothesis, it would be necessary to obtain longitudinal data by asking the T3 listeners to repeat the same task in a couple of years' time. The results of the T2 group suggest that, if the T3 listeners gain comparable English experience (both in quality and quantity) as the T2 listeners, the T3 group's discrimination accuracy for the English stimuli would improve. If, on the other hand, they become engaged in an occupation which does not require much use of English, their discrimination accuracy may be fossilized or even deteriorate (Flege & Liu 2001).

Furthermore, for future research, it would be informative to test Thai listeners with even less English experience than the T3 group, for example, primary (aged 6–11) or junior high-school (aged 12–14) students or adults who receive no or limited English input from native speakers in their daily life to assess how much benefit the L1-to-L2 phonological mapping gives them. It can be hypothesized here that these listeners would discriminate English stop contrasts less accurately than the T3 listeners if discrimination accuracy in L2 were more closely related to the specific experience in their L2 than to the phonological mapping between the L1 and L2. Alternatively, discrimination accuracy may be predicted by a combination of both of these elements. In other words, even the Thai listeners who have absolutely no experience with spoken English may still show a reasonably good level of discrimination accuracy for English stops as the benefit from the between-language mapping and abundant acoustic-phonetic information in the case of released stop tokens, but additional experience in English would further improve their cross-language perceptual performance.

Finally, it will be necessary to obtain empirical data on the perceived relationship between English and Thai final stops which would form the basis for different predictions. We are unaware of such perception data at present. It is possible that Thai stops are not mapped onto two separate English categories, but instead, onto multiple English categories with some degree of overlap. If such data were obtained, it might provide an alternative explanation for non-optimal discrimination accuracy by the AusE listeners in this study.

In sum, we have provided evidence for phonetic learning in L2 by some native Thai speakers. The more experience with L2, the better adjusted learners are to the specific phonetic realizations of L2 sounds. Our research demonstrated the need to take into account detailed acoustic phonetic information even for those phonemes which are shared between L1 and L2 in studying cross-language speech perception and L2 speech learning.

Acknowledgements

This research was partially supported by a University of Western Sydney seed grant. The original manuscript was prepared when the first author was affiliated with the Department of East Asian Languages and Literatures, University of Oregon, and the second author with the Department of Linguistics, Chulalongkorn University. Portions of this work were presented at the 15th Australian Language and Speech Conference held at Macquarie University in December 2005. We thank James

Flege for his guidance and Rhondda Fahey, Shunichi Ishihara, Joan Liang and Sue Spinks for subject recruitment and data collection, and Sorabud Rungrojsuwan and Colin Schoknecht for advice on the Thai speech materials. We also thank the SHLRC (Speech, Hearing and Language Research Centre) and the Department of Linguistics, Macquarie University, for the use of the perception laboratory and the Department of Linguistics, Faculty of Arts, Chulalongkorn University, for the use of the facilities for conducting the perception experiments in Thailand. We are indebted to John Esling, Adrian Simpson and three anonymous reviewers for valuable comments on earlier versions of our manuscript.

Appendix A: English test words used

Vowel	/p/	/t/	/k/	N
/iː/ (/i/)	peep$_{1, 3}$	Pete$_{2}$	peak$_{1, 2, 3}$	6
	beep$_{1, 3}$	beat$_{1, 2, 3}$	beak$_{1, 2}$	7
	seep$_{1, 2, 3}$	seat$_{1, 2, 3}$	seek$_{1, 2, 3}$	9
	–	meet$_{1, 2, 3}$	meek$_{3}$	4
	cheap$_{1, 2, 3}$	cheat$_{1, 2, 3}$	cheek$_{2}$	7
	weep$_{2}$	wheat$_{1, 3}$	week$_{1, 2, 3}$	6
/ɪ/ (/ɪ/)	sip$_{1, 3}$	sit$_{2, 3}$	sick$_{1, 2, 3}$	7
	lip$_{1}$	*not used*	lick$_{1, 2, 3}$	4
	hip$_{1, 2}$	hit$_{3}$	–	3
	–	kit$_{1}$	kick$_{2, 3}$	3
/æ/ (/æ/)	rap$_{1, 2a, 2b, 3}$	*not used*	rack$_{1}$	5
	–	bat$_{2}$	back$_{1, 3}$	3
	cap$_{1, 2}$	cat$_{1, 2, 3}$	–	5
	lap$_{1, 2a, 2b, 3}$	–	lack$_{3}$	5
/ʉː/ (/u/)	soup$_{1, 2}$	suit$_{3}$	–	3
/ɐː/ (/a/)	harp$_{1}$	heart$_{2, 3}$	–	3
	carp$_{3}$	cart$_{1, 2}$	–	3
	sharp$_{1, 2, 3}$	–	shark$_{1, 2, 3}$	6
/ɐ/ (/ʌ/)	cup$_{2}$	cut$_{1, 3}$	–	3
/əʉ/ or /əʊ/ (/oʊ/)	cope$_{1, 2, 3}$	coat$_{1}$	Coke$_{2, 3}$	6
/æɪ/ (/eɪ/)	–	fate$_{1, 2}$	fake$_{1, 2, 3}$	5
	–	late$_{1, 2}$	lake$_{3}$	3
	–	mate$_{1, 2, 3}$	*not used*	3
	tape$_{1, 2, 3}$	–	take$_{2, 3}$	5
/ɑe/ (/aɪ/)	–	bite$_{3}$	bike$_{1, 2}$	3
			Total:	117

The subscripts indicate different talkers. Phonetic symbols in the traditional system (Mitchell 1946) are in parentheses.

Appendix B: Thai test words used

Vowel	/p/	Gloss	/t/	Gloss	/k/	Gloss	N
/i/	จิบ$_{2,3}$ /cìp/*	to taste	จิต$_{1}$/cìt/	mind	จิก$_{1,2,3}$ /cìk/	to peck at	6
	จิ๊บ$_{1,2,3}$ /cíp/**	bird's sound	จิ๊ด$_{2}$ /cít/	small	จิ๊ก$_{1,3}$ /cík/	stole	6
/a/	จับ$_{2,3}$ /càp/	to catch	จัด$_{1,2,3}$ /càt/	hard	จัก$_{1}$ /càk/	to split	6
	ฉับ$_{2,3}$ /cʰàp/	quickly	ฉัตร$_{1,3}$ /cʰàt/	tiered umbrella	ฉัก$_{1}$ /cʰàk/	train's sound	5
	สับ$_{1,3}$ /sàp/	to chop	สัตว์$_{1,2}$ /sàt/	animal	สัก$_{2,3}$ /sàk/	to tattoo	6
	หับ$_{1,2,3}$ /hàp/	to close	*not used*	–	หัก$_{1,2}$ /hàk/	to break	5
	ปับ$_{1,2,3}$ /pàp/	immediate	ปัด$_{1,2,3}$ /pàt/	to brush away	*not used*	–	6
	ตับ$_{1,2}$ /tàp/	liver	ตัด$_{1,2,3}$ /tàt/	to cut	ตัก$_{3}$ /tàk/	to dip up	6
	กับ$_{1,2}$ /kàp/	with	กัด$_{1,2,3}$ /kàt/	to bite	กัก$_{3}$ /kàk/	to confine	6
	ผับ$_{1}$ /pʰàp/	thud	ผัด$_{1,2,3}$ /pʰàt/	to stir fry	ผัก$_{3}$ /pʰàk/	vegetable	5
	ดับ$_{2}$ /dàp/	to extinguish	ดัด$_{1,3}$ /dàt/	to bend	ดัก$_{1,2}$ /dàk/	to trap	5
	ซับ$_{1,3}$ /sáp/	to absorb	*not used*	–	ซัก$_{1,2,3}$ /sák/	to wash	5
	พับ$_{1,3}$ /pʰáp/	to fold	พัด$_{1,2}$ /pʰát/	to fan	พัก$_{2,3}$ /pʰák/	to rest	6
	ทับ$_{3}$ /tʰáp/	to put on top	ทัด$_{1,3}$ /tʰát/	equal	ทัก$_{1,2}$ /tʰák/	to greet	5
	คับ$_{2}$ /kʰáp/	tight	คัด$_{3}$ /kʰát/	to choose	คัก$_{1,2,3}$ /kʰák/	active	5
	นับ$_{1}$ /náp/	to count	นัด$_{2}$ /nát/	to make an appointment	นัก$_{1,2,3}$ /nák/	expert	5
	วับ$_{1,2,3}$ /wáp/	in a flash	วัด$_{1}$ /wát/	monastery	วัก$_{2,3}$ /wák/	to bail	6
/u/	จุบ$_{2,3}$ /cùp/	to draw	จุด$_{1}$ /cùt/	to ignite	จุก$_{1,2,3}$ /cùk/	stopper	6
	ฉุบ$_{1,2}$ /cʰùp/	paper, scissors, stone	ฉุด$_{2,3}$ /cʰùt/	to drag	ฉุก$_{1,3}$ /cʰùk/	to happen suddenly	6
	ปุบ$_{1,2}$ /pùp/	suddenly	ปุด$_{3}$ /pùt/	stitch	ปุก$_{1,2,3}$ /pùk/	to be fluffy	6
	บุบ$_{3}$ /bùp/	to grind	บุตร$_{1,2}$ /bùt/	child	บุก$_{1,2}$ /bùk/	to force	5
	not used	–	ชุด$_{1,2,3}$ /cʰút/	suite	*not used*	–	3
	ซุป$_{1,3}$ /súp/	soup	ทรุด$_{1,3}$ /sút/	worse	ซุก$_{2}$ /súk/	to snuggle	5
						Total:	125

* /`/ = a low tone; ** /´/ = a high tone
The subscripts indicate different talkers.

References

Abramson, Arthur S. 1972. Word-final stops in Thai. In Jimmy G. Harris & Richard B. Noss (eds.), *Tai phonetics and phonology*, 1–7. Central Institute of English Language, Office of State Universities, Faculty of Science, Mahidol University.

Abramson, Arthur S. 2001. The stability of distinctive vowel length in Thai. In M. R. Kalaya Tingsabadh & Arthur S. Abramson (eds.), *Essays in Tai linguistics*, 13–26. Bangkok: Chulalongkorn University Press.

Abramson, Arthur S. & Nianqi Ren. 1990. Distinctive vowel length: Duration vs. spectrum in Thai. *Journal of Phonetics* 18(2), 79–92.

Abramson, Arthur S. & M. R. Kalaya Tingsabadh. 1999. Thai final stops: Cross-language perception. *Phonetica* 56(3/4), 111–122.

Aoyama, Katsura, James E. Flege, Susan G. Guion, Reiko Akahane-Yamada & Tsuneo Yamada. 2004. Perceived phonetic dissimilarity and L2 speech learning: The case of Japanese /r/ and English /l/ and /r/. *Journal of Phonetics* 32(2), 233–250.

Bent, Tessa & Ann R. Bradlow. 2003. The interlanguage intelligibility benefit. *Journal of the Acoustical Society of America* 114(3), 1600–1610.

Best, Catherine T. 1995. A direct realist view of cross-language speech. In Winifred Strange (ed.), *Speech perception and linguistic experience: Issues in cross-language research*, 171–204. Timonium, MD: York Press.

Best, Catherine T. & Winifred Strange. 1992. Effects of phonological and phonetic factors on cross-language perception of approximants. *Journal of Phonetics* 20(3), 305–330.

Bond, Z. S. & Thomas J. Moore. 1994. A note on the acoustic-phonetic characteristics of inadvertently clear speech. *Speech Communication* 14(4), 324–337.

Broersma, Mirjam. 2005. Perception of familiar contrasts in unfamiliar positions. *Journal of the Acoustical Society of America* 117(6), 3890–3901.

Byrd, Dani. 1992. Preliminary results on speaker-dependent variation in the TIMIT database. *Journal of the Acoustical Society of America* 92(1), 593–596.

Byrd, Dani. 1993. 54,000 American stops. *UCLA Working Papers in Phonetics* 83, 97–115.

Byrd, Dani. 1994. Relations of sex and dialect on reduction. *Speech Communication* 15(1/2), 39–54.

Cho, Taehong & James M. McQueen. 2006. Phonological versus phonetic cues in native and non-native listening: Korean and Dutch listeners' perception of Dutch and English consonants. *Journal of the Acoustical Society of America* 119(5), 3085–3096.

Ciocca, Valter, Livia Wong & Lydia So. 1994. An acoustic analysis of unreleased stop consonants in word-final position. *3rd International Conference on Spoken Language Processing*, Yokohama, Japan, 1131–1134.

Cochrane, G. R. 1970. Some vowel durations in Australian English. *Phonetica* 22, 240–250.

Cox, Felicity. 2006a. Australian English pronunciation into the 21st century. *Prospect* 21(1), 3–21.

Cox, Felicity. 2006b. The acoustic characteristics of /hVd/ vowels in the speech of some Australian teenagers. *Australian Journal of Linguistics* 26(2), 147–179.

Crystal, Thomas H. & Arthur S. House. 1988. Segmental duration in connected-speech signals: Current results. *Journal of the Acoustical Society of America* 83(4), 1553–1573.

Flege, James E. 1987. The production of 'new' and 'similar' phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* 15(1), 47–65.

Flege, James E. 1988. Factors affecting degree of perceived foreign accent in English sentences. *Journal of the Acoustical Society of America* 84(1), 70–79.

Flege, James E. 2003. Methods for assessing the perception of vowels in a second language. In Elisabetta Fava & Alberto Mioni (eds.), *Issues in clinical linguistics*, 3–28. Padova: UniPress.

Flege, James E., David Birdsong, Ellen Bialystok, Molly Mack, Hyekyung Sung & Kimiko Tsukada. 2006. Degree of foreign accent in English sentences produced by Korean children and adults. *Journal of Phonetics* 34(2), 153–175.

Flege, James E., Ocke-Schwen Bohn & Sungyoung Jang. 1997. Effects of experience on non-native speakers' production and perception of English vowels. *Journal of Phonetics* 25(4), 437–470.

Flege, James E. & Kathryn L. Fletcher. 1992. Talker and listener effects on degree of perceived foreign accent. *Journal of the Acoustical Society of America* 91(1), 370–389.

Flege, James E. & Serena Liu. 2001. The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition* 23(4), 527–552.

Flege, James E. & Ian R. A. MacKay. 2004. Perceiving vowels in a second language. *Studies in Second Language Acquisition* 26(1), 1–34.

Flege, James E., Ian R. A. MacKay & Diane Meador. 1999. Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America* 106(5), 2973–2987.

Flege, James E., Naoyuki Takagi & Virginia Mann. 1996. Lexical familiarity and English-language experience affect Japanese adults' perception of /ɹ/ and /l/. *Journal of the Acoustical Society of America* 99(2), 1161–1173.

Gandour, Jack. 1984. Vowel duration in Thai. *Crossroads: An Interdisciplinary Journal of Southeast Asian Studies* 2(1), 59–64.

Guion, Susan G. 2003. The vowel systems of Quichua–Spanish bilinguals: Age of acquisition effects on the mutual influence of the first and second languages. *Phonetica* 60(2), 98–128.

Guion, Susan G., James E. Flege, Reiko Akahane-Yamada & Jesica C. Pruitt. 2000a. 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.

Guion, Susan G., James E. Flege & Jonathan D. Loftin. 2000b. The effect of L1 use on pronunciation in Quichua-Spanish bilinguals. *Journal of Phonetics* 28(1), 27–42.

Halle, Morris, George W. Hughes & Jean-Pierre A. Radley. 1957. Acoustic properties of stop consonants. *Journal of the Acoustical Society of America* 29(1), 107–116.

Hallé, Pierre A., Catherine T. Best & Andrea Levitt. 1999. Phonetic vs. phonological influences on French listeners' perception of American English approximants. *Journal of Phonetics* 27(3), 281–306.

Hancin-Bhatt, Barbara. 2000. Optimality in second language phonology: Codas in Thai ESL. *Second Language Research* 16(3), 201–232.

Harrington, Jonathan, Felicity Cox & Zoe Evans. 1997. An acoustic phonetic study of broad, general, and cultivated Australian English vowels. *Australian Journal of Linguistics* 17(2), 155–184.

Householder, Fred W. Jr. 1956. Unreleased PTK in American English. In Morris Halle, Horace G. Lunt, Hugh McLean & Cornelis H. van Schooneveld. (eds.), *For Roman Jacobson*, 235–244. The Hague: Mouton.

Kenstowicz, Michael & Atiwong Suchato. 2006. Issues in loanword adaptation: A case study from Thai. *Lingua* 116(7), 921–949.

Ladefoged, Peter. 2006. *A course in phonetics*. Boston, MA: Thomson/Wadsworth.

Lado, Robert. 1957. *Linguistics across cultures*. Ann Arbor, MI: The University of Michigan Press.

Lambacher, Stephen, William L. Martens, Brian Nelson & Jonathan Berman. 2001. Identification of English fricatives by Japanese listeners: The influence of vowel context on sensitivity and response bias. *Acoustical Science and Technology* 22(5), 334–343.

Lisker, Leigh. 1999. Perceiving final voiceless stops without release: Effects of preceding monophthongs versus nonmonophthongs. *Phonetica* 56(1/2), 44–55.

MacKay, Ian R. A., James E. Flege, Thorsten Piske & Carlo Schirru. 2001. Category restructuring during second-language (L2) speech acquisition. *Journal of the Acoustical Society of America* 110(1), 516–528.

Malécot, André. 1958. The role of releases in the identification of released final stops. *Language* 34(3), 370–380.

Mitchell, Alexander G. 1946. *The pronunciation of English in Australia*. Sydney: Angus and Robertson.

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

Parker, Frank & Thomas Walsh. 1981. Voicing cues as a function of tense/lax distinction in vowels. *Journal of Phonetics* 9, 353–358.

Picheny, Michael A., Nathaniel I. Durlach & Louis D. Braida. 1985. Speaking clearly for the hard of hearing I: Intelligibility differences between clear and conversational speech. *Journal of Speech, Language and Hearing Research* 28(1), 96–103.

Picheny, Michael A., Nathaniel I. Durlach & Louis D. Braida. 1986. Speaking clearly for the hard of hearing II: Acoustic characteristics of clear and conversational speech. *Journal of Speech, Language and Hearing Research* 29(4), 434–446.

Piske, Thorsten, James E. Flege & Ian R. A. MacKay. 2001. Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29(2), 191–215.

Sancier, Michele C. & Carol A. Fowler. 1997. Gestural drift in a bilingual speaker of Brazilian Portuguese. *Journal of Phonetics* 25(4), 421–436.

Snodgrass, Joan G., Gail Levy-Berger & Martin Haydon. 1985. *Human Experimental Psychology*. New York: Oxford University Press.

Tingsabadh, M. R. Kalaya & Arthur S. Abramson. 1993. Thai. *Journal of the International Phonetic Association* 23(1), 24–28.

Tsukada, Kimiko. 2006. Cross-language perception of final stops in Thai and English. *Bilingualism: Language and Cognition* 9(3), 309–318.

Tsukada, Kimiko, David Birdsong, Ellen Bialystok, Molly Mack, Hyekyung Sung & James E. Flege. 2005. A developmental study of English vowel production and perception by native Korean adults and children. *Journal of Phonetics* 33(3), 263–290.

Tsukada, Kimiko, David Birdsong, Molly Mack, Hyekyung Sung, Ellen Bialystok & James E. Flege. 2004. Release bursts in English word-final voiceless stops produced by native English and Korean adults and children. *Phonetica* 61(2/3), 67–83.

Tsukada, Kimiko & Shunichi Ishihara. 2007. The effect of first language (L1) in cross-language speech perception: Comparison of word-final stop discrimination by English, Japanese and Thai listeners. *Journal of the Phonetic Society of Japan* 11(1), 82–92.

Wang, William S-Y. 1959. Transition and release as perceptual cues for final plosives. *Journal of Speech Language and Hearing Research* 2(1), 66–73.

Watson, Catherine I. & Jonathan Harrington. 1999. Acoustic evidence for dynamic formant trajectories in Australian English vowels. *Journal of the Acoustical Society of America* 106(1), 458–468.

Wayland, Ratree & Susan Guion. 2003. Perceptual discrimination of Thai tones by naïve and experienced learners of Thai. *Applied Psycholinguistics* 24(1), 113–129.

Wayland, Ratree P. & Susan G. Guion. 2004. Training English and Chinese listeners to perceive Thai tones: A preliminary report. *Language Learning* 54(4), 681–712.