Perception of Central Thai Tones and Segments by Thai and Australian Adults

Denis Burnham
Kathryn Kirkwood
School of Psychology
University of NSW
Sydney Australia

Sudaporn Luksaneeyanawin
Supatra Pansottee
Linguistics Research Unit
Chulalongkorn University
Bangkok Thailand

Introduction Thai and Australian adults were tested for their ability to discriminate pairs of bilabial stop consonant plus /a:/ vowel syllables spoken by a Thai speaker. Three different speech contrasts were incorporated into the design of the experiment: (i) prevoiced bilabial stops vs voiceless aspirated bilabial stops - /b/ vs /pʰ/; (ii) prevoiced bilabial stops vs voiced unaspirated bilabial stops - /b/ vs /p/; and (iii) tonal contrasts incorporating all possible combinations of the five Central Thai tones. The three main purposes of the study were (a) to investigate the relative salience of Thai phonological distinctions, (b) to study the effects of linguistic experience on speech perception, and (c) to investigate hemispheric specialization in speech perception. Details of these are set out below.

(a) Relative Salience of Thai Contrasts
(i) Tonal Contrasts: Central Thai has five tones (mid, high, low, rising, and falling). Many studies have investigated the relative perceptual salience of these for Thais. For example, Gandour (1979) conducted a multidimensional scaling analysis on Thais' discrimination of various pitch contours and found three tone-related dimensions. The relatively low-level auditory dimension of average pitch was the most important factor in subjects' perceptual judgements, and is also an important cue for speakers of non-tonal languages and for speakers of tonal languages other than Thai (Gandour, 1983). The other two dimensions were the direction dimension, which served to distinguish between rising and falling contours on the basis of pitch movement rather than absolute start and end points, and the slope dimension, which served to distinguish between what Abramson (1978) has called static (mid, low, high) and
dynamic (rising, falling) tones. These results agree to some extent with Abramson's (1978, 1986) studies with Thais in which he found that differences in pitch levels are sufficient for the identification of static tones, although this is enhanced by slow $f_0$ movement.

(ii) Consonantal Contrasts: The bilabial stop voiced-voiceless contrast is phonologically relevant in English and understandably this is discriminated more easily by English speakers than the irrelevant prevoiced-voiced contrast. However, there is some indication that this is not just a function of linguistic experience. It has been found that 6- and 10-month-old infants also have greater difficulty perceiving the prevoiced-voiced contrast (Aslin, et al., 1981; Burnham, et al., 1991), even though at this age this bias is unlikely to be linguistically based (Burnham, 1986). Results of a study by Pisoni (1977) suggest that there may be a psychoacoustic element to this bias. Investigation of the perception of these two contrasts by Thai and English speakers in this study will provide a good test of this hypothesis.

(iii) Consonantal vs Tonal Contrasts: Some information about the relative salience of Thai consonantal and tonal contrasts may be gleaned from developmental studies. Clumeck (1980) reports that the onset of the lexical use of tonal and segmental distinctions coincide, each of these first emerging around 11 months. However, acquisition of tonal distinctions appears to be easier, because these are completely acquired by around 23 months before the acquisition of segments is completed (Tuaycharoen, 1977; Luksaneeyanawin, 1976; Clumeck, 1980). There is also evidence from a 10- to 11-month-old child that tones are more perceptually salient than segments (Li & Thompson, 1977). These data suggest that tone differences are more basic than segmental differences and should be more easily discriminated by English speakers when phonological relevance is equated.

(b) The Effects of Linguistic Experience Many studies have shown that linguistic experience systematically biases speech perception abilities towards the phonological distinctions present in the ambient language (Aslin et al., 1981; Burnham et al., 1991; Werker & Tees, 1983, 1984), such that adults have more restricted perceptual abilities than infants in the same linguistic environment. Werker concludes
that this shift from infancy to adulthood does not entail sensori-neural loss because under certain conditions adults can still discriminate phonologically-irrelevant contrasts (Werker & Logan, 1985; Werker & Tees, 1984). There also seems to be an attentional shift associated with the onset of reading which biases children towards the perception of phonemes which are relevant in the ambient language (Burnham, 1986). The effects of linguistic experience were studied in the current study by the comparison of Thai and English speakers' perception of relevant and irrelevant distinctions in a 500msec interstimulus interval (ISI) condition and a 1500msec ISI condition. These two intervals have been shown by Werker & Logan (1985) to induce a language-general phonetic mode of perception and a language-specific phonological mode of perception, respectively. It was expected that native English speakers' performance on the irrelevant consonantal contrast and the tonal contrasts should be better at 500 than at 1500 msec.

(c) Hemispheric Specialization for Speech The left hemisphere of the human brain is said to be specialized for language processing with between 93% and 100% of right-handers being left hemisphere dominant for language (Hiscock & Kinsbourne, 1980; Kolb & Whishaw, 1990). Behavioural studies of this specialization have used the dichotic task, in which different auditory inputs are delivered to each ear and the subject is required to identify input to one particular ear. Using this procedure Haggard and Parkinson (1971) found stronger right ear advantages (REAs) for place than voicing contrasts, and Studdert-Kennedy and Shankweiler (1967) found stronger REAs for voiceless place contrasts than for voiced place contrasts. Thus phonetic cues seem to vary in their potency in the production of ear advantages.

With regard to tone, speakers of languages which do not use tone at a lexical level either show a left ear advantage (LEA) or no ear advantage for tonal identifications and discriminations. For example, Murray (1986) found a significant REA for speech syllables and an LEA for simple tones of different frequencies. On the other hand, speakers of tonal languages show REAs for tonal contrasts. Van Lancker and Fromkin (1973) found an REA for Thai speakers' identification of dichotically presented consonantal and tonal contrasts but no ear advantage for hummed versions of the five Thai tones, while English
speakers showed an REA for consonant differences but no ear advantage with tonal contrasts or hums.

From this behavioural evidence it appears that speakers of tonal languages process tone in the left hemisphere along with consonantal distinctions. In this study it was expected that Thai speakers should show REAs for all three contrasts while English speakers should show an REA for the /b/-/pʰ/ contrast, possibly for the /b/-/p/ contrast, but certainly not for the tonal contrasts.

Method
Subjects 48 right-handed adult subjects were tested. Of these 24 were native Australian English speakers with little or no experience with other languages. The other 24 were native Thai speakers, who could also speak English with varying degrees of proficiency.

Design A native language (Thai/English) x ISI (500msec/1500msec) x Contrast (/b/-/pʰ/, /b/-/p/, tonal contrasts) x Ear of Presentation (Left/Right) design was employed. All subjects were tested on all three contrasts with both ears of presentation. Half of the subjects in each language group were tested with an ISI of 500msec between the two stimuli to be discriminated and half were tested with the 1500msec ISI condition. In each language by ISI subgroup half the subjects were males and half were females. Presentation order of the three blocks of contrasts was balanced between subjects.

Stimulus Materials Speech tokens were all produced by a native Thai female (S.L. in the author list). Bilabial stop consonant plus /a:/ vowel tokens, /ba:/:, /pa:/ and /pʰa:/, were used. For the consonantal contrasts these were all presented with mid tone. For the tone contrasts the prevoiced syllable, /ba:/, was used to carry the five tones. For each of the seven different consonant-vowel pairs, five exemplars were produced by the speaker. These were digitized and stored on disk of the laboratory computer.

Three different Thai contrasts were tested in the experiment: /ba:/ vs /pʰa:/, which is phonologically relevant in English; /ba:/ vs /pa:/, which is phonologically irrelevant in English; and tone contrasts, e.g., /báː/ vs /báː/ which again are irrelevant in English. All possible pairings of the five tones were used, a total of 10 possible tone