Perceptual factors in the acquisition of FL vowels:
The locations of vowel prototypes

Satsuki Nakai
University of Edinburgh
(Dept. of Linguistics, 40 George Sq., Edinburgh EH8 9LL, UK, satsuki@ling.ed.ac.uk)

Kuhl (1991) proposed that phonetic prototypes assimilate other members of the category, which results in poor discrimination sensitivity around the prototypes (perceptual magnet effect). According to Kuhl (1993), perceptual magnet effect interferes with adults' acquisition of FL vowels, by pulling the target vowels towards the L1 prototype. The validity of Kuhl's 1991 experiment, however, has been questioned by a number of researchers (e.g., Lively, 1993) who replicated her study and obtained results that are discrepant from Kuhl (1991). In this paper, an attempt was made to answer one such question raised by those researchers, i.e., the locations of vowel prototypes in relation to their production means. The implications for the acquisition of foreign-language (FL) vowels will be discussed.

1. Locations of Vowel Prototypes

The notion of phonetic prototypes as mental representations of phonetic categories has been widely accepted in the field of speech perception. Phonetic prototypes are perceived as better instances of the category (Grieser and Kuhl, 1989; Miller, 1994), and have potency over other members of the category, as shown in selective adaptation experiments (Miller, et al., 1983; Samuel 1982) and dichotic competition experiments (Repp, 1977).

In Kuhl's (1991) further investigation of phonetic prototypes using a range of /i/ stimuli, the stimulus that received the highest rating in a goodness rating task has formant-frequency values that coincide with the average values obtained from male speakers by Peterson and Barney (1952). Lively (1993), however, reports that in a goodness rating task using an identical stimulus set as in Kuhl (1991), listeners chose more extreme /i/ as the best tokens. Sussman and Lauckner-Morano (1995) and Lotto, et al. (1996) also report that listeners preferred more extreme stimuli than Kuhl (1991) prototype. Indication of listeners' preferences for extreme vowels can be also found in other papers such as Johnson, et al. (1993) and Bradlow (1993).

Thus, evidence seems to suggest that vowel prototypes may be more extreme than production average. However, caution must be used in interpreting the results of the above studies, for they all use synthesised, steady-state vowel stimuli for perception, while vowels in production data are in consonantal contexts, which are bound to be less extreme than their isolated versions due to coarticulation. In other words, the materials in the production tasks and stimuli in the perception tasks are not strictly comparable.

2. Experiments and Results

The chief aim of Experiments 1a and 1b is to re-examine the relationships between the locations of vowel prototypes in the vowel space and production means, with comparable materials for perception and production tasks. The Greek and Japanese vowels were investigated for comparison, for the two languages have phonologically identical sets of vowels.

2.1 Experiment 1a

In Experiment 1a, the production average of the Greek and Japanese vowels /a/, /i/, /u/ (in the case of Japanese, [ɯ]), /e/ and /o/ were obtained. Five Greek and 5

Japanese adult male native speakers participated in this experiment. Vowels were produced in isolation so that the data would be comparable with that from Experiment 1b, where synthesised steady-state vowels were used for a perception task. Speakers were instructed to produce the vowels in a natural manner. Recording was conducted in studio conditions. Each vowel was transcribed in Greek and Japanese orthography on 5 separate cards and presented to the subjects in a random order. Thus, 25 vowels were produced by each speaker. Both LPC spectra and spectrograms were used to measure F1 and F2 of the steady portion of each vowel.

2.2 Experiment 1b

In Experiment 1b, Johnson et al.'s (1993) MOA (the method of adjustment) procedure was adopted to locate Greek and Japanese vowel prototypes. In MOA, subjects can adjust parameters of a speech synthesiser till the synthesiser produces the sound they have in mind.

Five Greek subjects and 4 Japanese subjects from Experiment 1a[1] were instructed to choose stimuli that were 'the closest' to their L1 vowels, from a 15 x 25 grid displayed on a computer screen. The grid consisted of 339 synthesised steady-state vowels systematically varying in F1 and F2. The range of F1 and F2 is determined so that it would cover the production range of the Greek and Japanese vowels obtained in a pilot experiment, and slightly differs from Johnson, et al. (1993). F1 varied from 214Hz to 841Hz in 15 steps along the vertical side of the grid, while F2 varied from 657Hz to 2260Hz in 25 steps along the horizontal side. The step size was approximately .38Bark for F1 and .34Bark for F2, slightly larger than the just-noticeable differences for vowel formants reported in Flanagan (1957). In the corner of the grid where F1 and F2 were close to each other, it was visually displayed that no sounds could be produced. The cut-off point was where the difference between F1 and F2 was smaller than 200Hz.

Stimuli assigned to each grid were produced using Sensimetrics SenSyn speech synthesiser software (Klatt, 1980). Each stimulus had F1 and F2 designated by its location in the grid. The stimuli were 250ms long, with 220ms of fixed amplitude that gradually attenuated in the last 30ms. F0 was kept constant at 120Hz during the first half and then was decreased gradually to 105Hz during the second half. F3 was estimated using regression formulas in Nearey (1989). F4 was set at least 300Hz higher than F3 and not lower than 3500Hz. The bandwidths of F1, F2 and F3 were calculated using similar regression formulas (see Johnson et al., 1993, p510). The bandwidth of F4 was kept at 200Hz.

Five repetitions of the 5 vowels (/a/, /i/, /u/ or [ɯ], /e/ and /o/) were presented above the grid on the computer screen in a random order, yielding 25 responses from each subject.

2.3 Results

Even when compared with the production of more comparable, isolated vowels, both Greek and Japanese subjects seem to perceive vowels that are more extreme as 'the closest' to their L1 vowels, though not in all cases, as shown in Figs. 1-7 below. No preference for extremity was observed in the choice of the low, central vowel /a/ (see Fig. 3). Furthermore, preference for extremity was observed only along F1 dimension in the case of the high, back vowel /u/ (see Figs. 2 and 5).

An ANOVA was run on F1 and F2 of the above vowels with condition (production and MOA) and subject as independent factors, and the real effect of condition was calculated by means of repeated measures. As for F1 of /i/ and /u/, there was a significant effect of condition; $F(1,8) = 10.9$, $p<.025$, and $F(1,8) = 22.4$, $p<.01$, respectively. In the case of F1 of

[1] One Japanese speaker who participated in Experiment 1a was not available for Experiment 1b.

/a/, there is no significant effect of condition; p=.74. As for F2 of /i/, the effect of condition was not significant; $F(1,8) = 7.04$, $p>.05$. This seems to be due to the ceiling effect in MOA task. The maximum F2 in MOA task was 2260Hz, and 6 out of 9 subjects consistently chose F2 value of 2260Hz for /i/, regardless of F2 value in production. As for F2 of /u/, the main effect of condition was not significant; $p = .08$. The effect of condition on F2 of /e/ and /o/ was significant; $F(1,8) = 31.0$, $p<.001$, and $F(1,8) = 26.1$, $p<.001$, respectively.

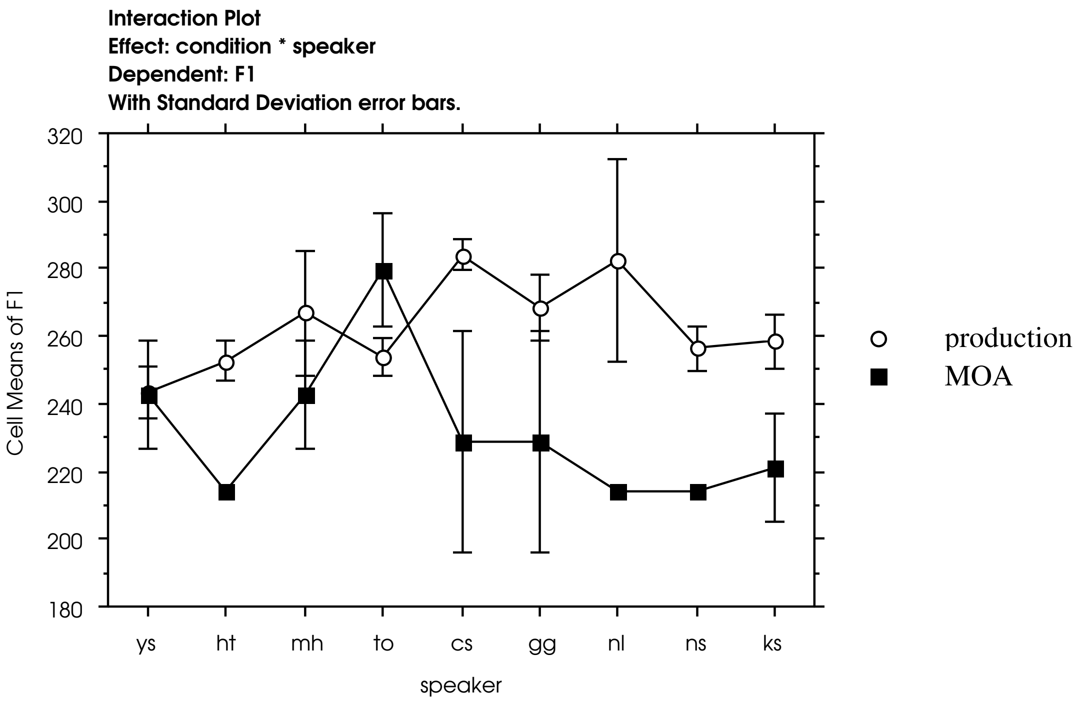

Figure 1 : F1 of /i/, Production Average & MOA with Standard Deviation

(ys, ht, mh, to = Japanese subjects; cs, gg, nl, ns, ks = Greek subjects)

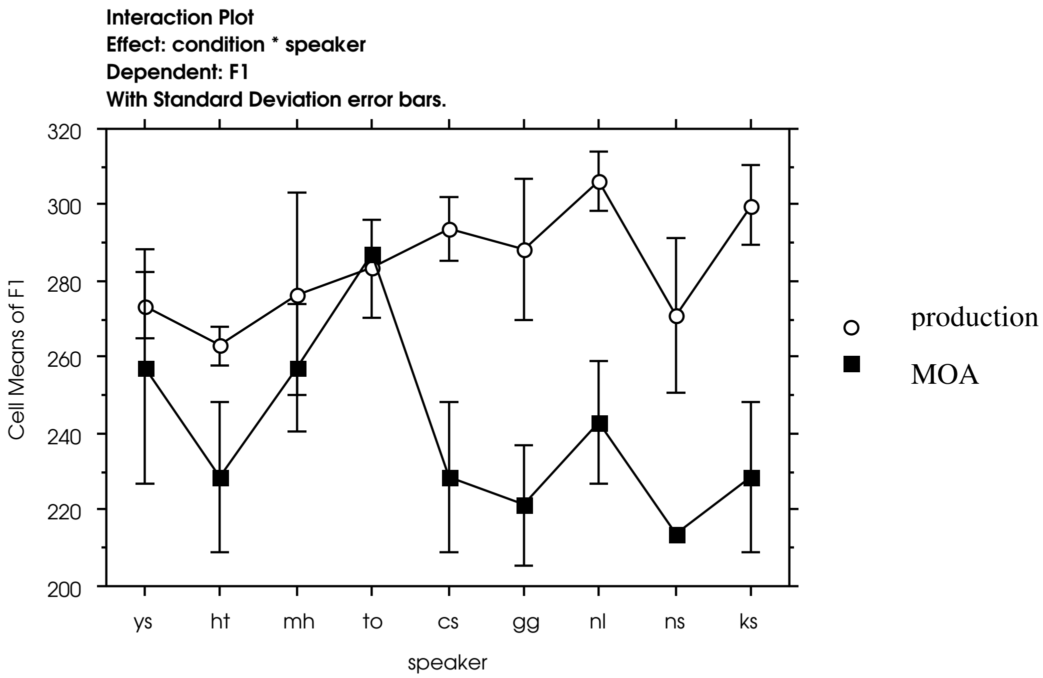

Figure 2 : F1 of /u/, Production Average & MOA with Standard Deviation

(ys, ht, mh, to = Japanese subjects; cs, gg, nl, ns, ks = Greek subjects)

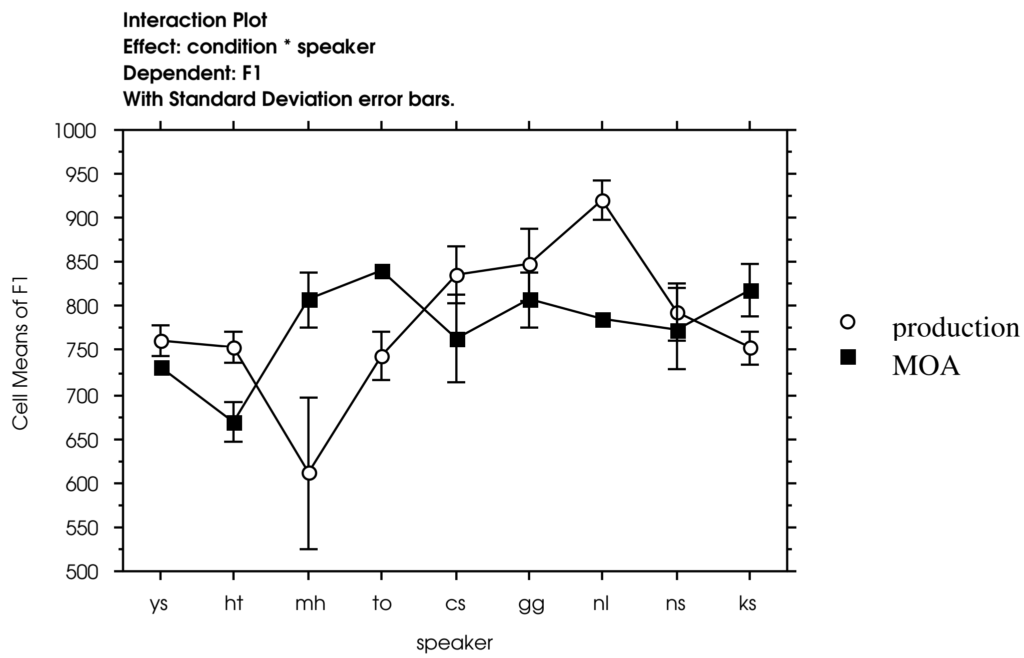

Figure 3 : F1 of /a/, Production Average & MOA with Standard Deviation

(ys, ht, mh, to = Japanese subjects; cs, gg, nl, ns, ks = Greek subjects)

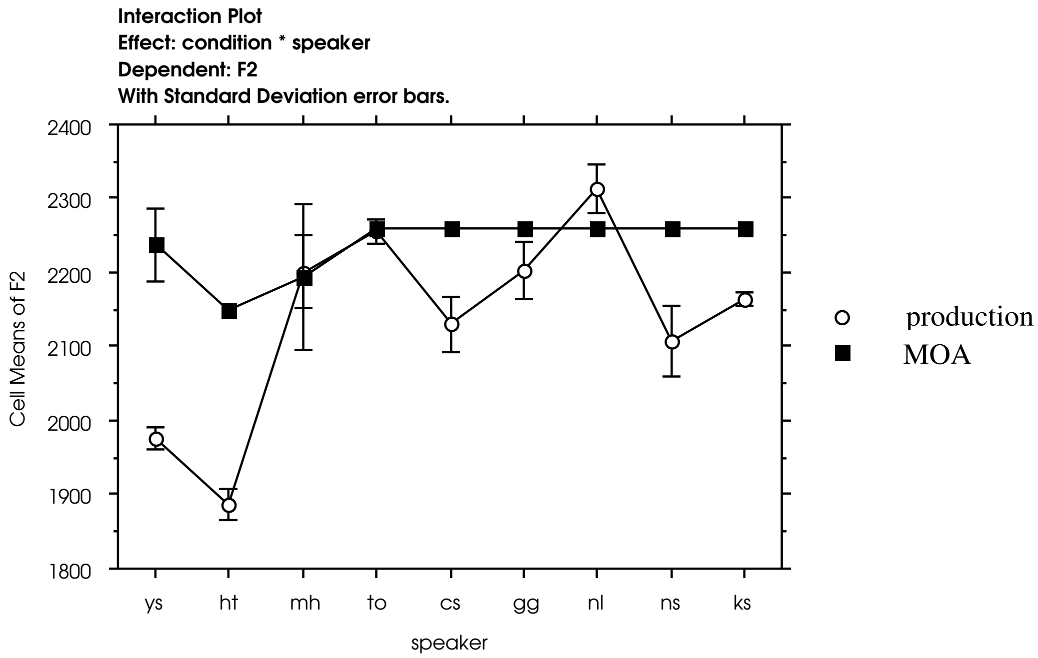

Figure 4 : F2 of /i/, Production Average & MOA with Standard Deviation

(ys, ht, mh, to = Japanese subjects; cs, gg, nl, ns, ks = Greek subjects)

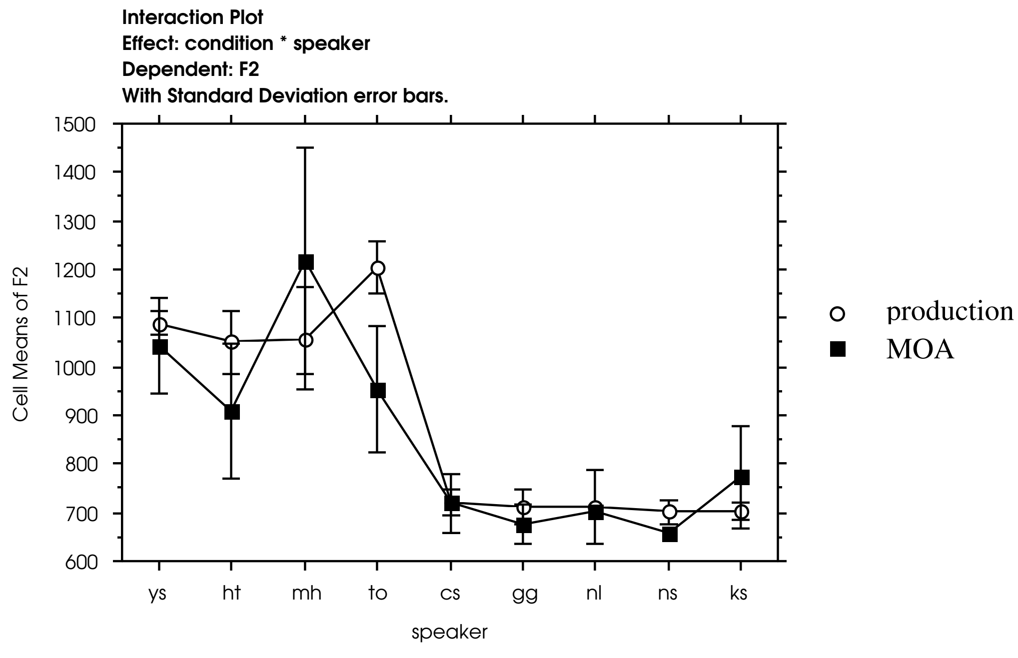

Figure 5 : F2 of /u/, Production Average & MOA with Standard Deviation

(ys, ht, mh, to = Japanese subjects; cs, gg, nl, ns, ks = Greek subjects)

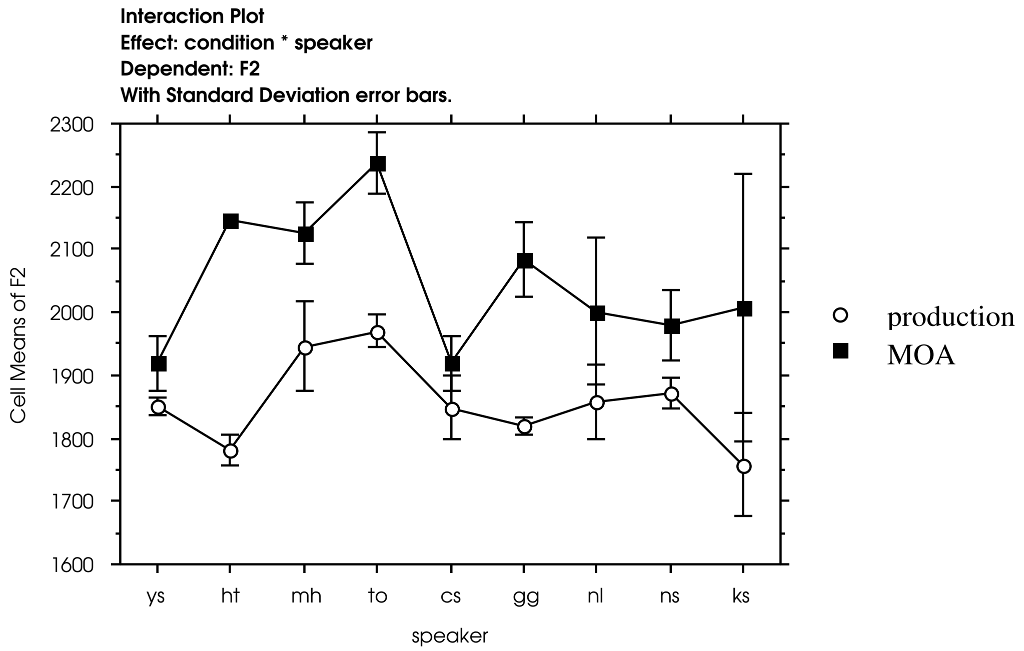

Figure 6 : F2 of /e/, Production Average & MOA with Standard Deviation

(ys, ht, mh, to = Japanese subjects; cs, gg, nl, ns, ks = Greek subjects)

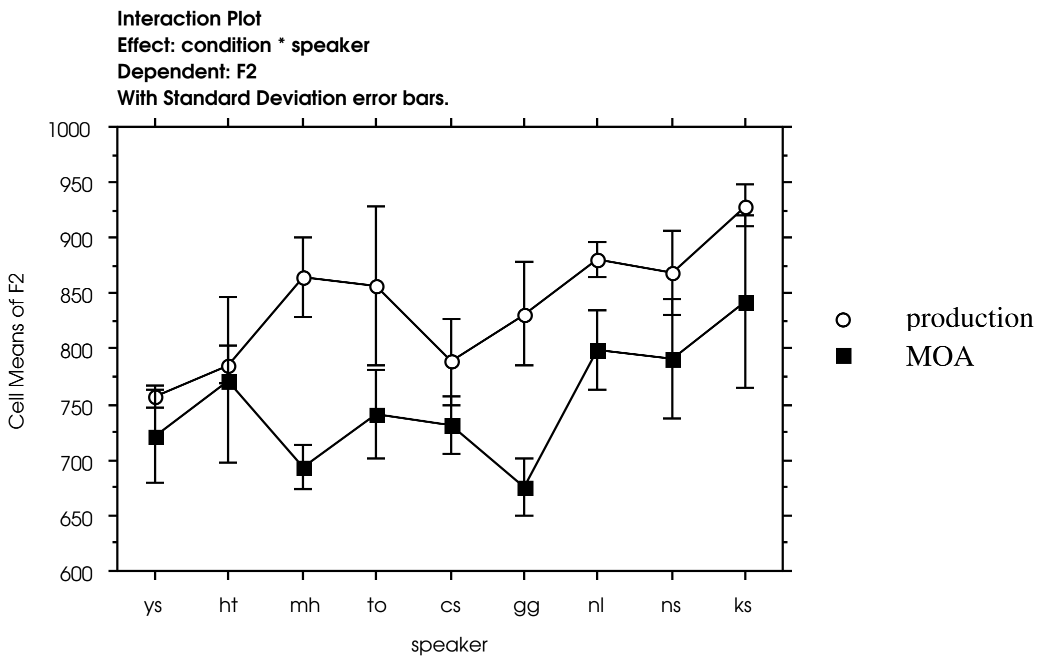

Figure 7 : F2 of /o/, Production Average & MOA with Standard Deviation

(ys, ht, mh, to = Japanese subjects; cs, gg, nl, ns, ks = Greek subjects)

3. Implications for the Acquisition of FL Vowels

The results of Experiments 1a and 1b suggest that Greek and Japanese listeners prefer more extreme vowels to production means for the high, front and back vowels /i/ and /u/ and the mid, front and back vowels /e/ and /o/, but not for the low central vowel /a/. Furthermore, the results suggest that preference for extreme vowels may not be exhibited in a uniform fashion. Subjects in this study exhibited preferences for extremity both along F1 and F2 dimension with the high, front vowel /i/, but only along F1 with the high, back vowel /u/.

According to Kuhl (1991), it is more difficult to discriminate between a prototypical /i/ and its neighbours in the acoustic vowel space than it is between a non-prototypical /i/ and its neighbours, provided that the psychoacoutsic distance between the stimuli is kept constant. Kuhl conceptualises the results as a ‘perceptual magnet effect’, for it appears as if the prototype perceptually assimilated other category members, causing the perceptual vowel space to shrink towards itself. According to Kuhl’s (1993) Native-Language Magnet (NLM) theory, vowel prototypes interfere with adults’ acquisition of FL vowels by assimilating their neighbours.

The validity of Kuhl (1991) has been questioned by a number of researchers, who replicated her study and obtained discrepant results from the original study. As mentioned earlier, unlike in Kuhl (1991), listeners preferred /i/ stimuli that are more extreme than production means in later studies, and their discrimination sensitivities were found to be low around these more extreme stimuli (see Lively, 1993; Kuhl and Iverson, 1995; Sussman and Lauckner-Morano, 1995). Considering evidence suggesting that vowel extremity may correlate with poor discrimination sensitivity (e.g., Schouten and van Hessen, 1992; Bohn and Polka, 1997), it is not clear whether poorer discrimination sensitivity observed around the most preferred /i/ stimuli in the above studies is due to the prototypicality of the stimuli or their extremity.

The question of whether it is vowel extremity or prototypicality that correlates with poor discrimination sensitivity is an interesting issue from the viewpoint of FL acquisition. Poor discrimination sensitivity signifies similarity, which, in accordance with current models of FL phonetics (e.g., Flege, 1987), leads to difficulty in acquisition. If vowel extremity correlates with poor discrimination sensitivity, there may be directional asymmetries in the ease of acquisition of FL vowels. Depending on whether a similar L1 vowel is more or less extreme than a target FL vowel, the perceptual distance between the two may differ, even if the psychoacoustic distance between the two is the same. On the other hand, if it is prototypicality that lowers discrimination sensitivity, there will be no such directional asymmetry in the ease of acquiring FL vowels.

References

- Bohn, O.-S. & Polka, L. (1997). ‘The nature and significance of perceptual biases in infant vowel perception’, *Paper presented at GALA 97*, Edinburgh Univ.
- Bradlow, A. R. (1993). *Language-specific and universal aspects of vowel production and perception: a cross-linguistic study of vowel inventories,* DMLL Publications: Ithaca.
- Flanagan, J. (1957). ‘Estimates of the maximum precision necessary in quantizing certain “dimensions” of vowel sounds’ *JASA, 29*, 533-4.
- Flege, J. E. (1987). ‘Effects of equivalance classification on the production of foreign language speech sounds’ in James, A. & Leather, J. (eds.) *Sound Patterns in Second Language Acquisition,* 9-40.
- Grieser, D. & Kuhl, P. K. (1989). ‘Categorization of speech by infants: Support for speech-sound prototypes’ *Developmental Psychology, 25 (4)*, 577-588.
- Johnson, K., et al. (1993). ‘The hyperspace effect: Phonetic targets are

hyperarticulated' *Language, 69 (3),* 505-528.
- Klatt, D. (1980). 'Software for a cascade/parallel formant synthesizer' *JASA, 67,* 971-95.
- Kuhl, P. K. (1991). 'Human adults and human infants exhibit a "perceptual magnet effect' for the prototypes of speech sounds, moneys do not' *Perception and Psychophysics, 50 (2),* 93-107.
- Kuhl, P. K. (1993). 'Early linguistic experience and phonetic perception: implications for theories of developmental speech perception' *Journal of Phonetics, 21,* 125-139.
- Lively, S. (1993). 'An examination of the "perceptual magnet" effect' *JASA 93*, 2423.
- Lotto, A. J. , et al. (1996). 'Effects of language experience on perceptual organization of vowel sounds' *Paper presented at the 5th conference on Laboratory phonology,* Northwestern Univ., Evanston, ILL.
- Miller, J. L., et al. (1983). 'A possible auditory basis for internal structure of phonetic categories' *JASA, 73*, 2124-2133.
- Miller, J. L. (1994). 'On the internal structure of phonetic categories: a progress report' *Cognition 50,* 271-285.
- Nearey, T. M. (1989). 'Static, dynamic and relational properties in vowel perception' *JASA, 85,* 2088-113.
- Peterson, G. & Barney, H. (1952). 'Control methods used in a study of the vowels' *JASA, 24 (2)*, 175-184.
- Repp, B. H. (1977). 'Dichotic competition of speech sounds: The role of acoustic stimulus structure' *J. of Experimental Psychology: Human Perception & Performance, 3*, 37-50.
- Samuel, A.G. (1982) 'Phonetic prototypes' *Perception & Psychophysics, 31,* 307-314.
- Schouten, M. E. H. and van Hessen, A. J. (1992). 'Modeling phoneme perception I: categorical perception' *JASA, 92(4),* 539-552.
- Sussman, J. E. & Lauckner-Morano V. J. (1995). 'Further tests of the "perceptual magnet effect" in the perception of [i]: identification and change/no-change discrimination' *JASA, 97 (1),* 539-552.