SIMULTANEOUS BILINGUALS AND FLEGE'S SPEECH LEARNING MODEL

Ian Watson
University of Oxford

ian.watson@chch.ox.ac.uk

ABSTRACT

The applicability of Flege's Speech Learning Model (SLM) to simultaneous bilinguals is examined in related experiments on the acquisition of the production and perception of the voicing contrast in simultaneous French-English bilinguals. The results show that the SLM can account for the data providing account is taken of the degree of bilinguals' exposure to each language.

Keywords: Simultaneous Bilingual, Acquisition.

1. INTRODUCTION

We have some data on the effects of simultaneous bilingualism on speech production or perception (e.g. [1,6,]), but as yet no model of these effects has been proposed to match models of L2 acquisition, notably Flege's Speech Learning Model (henceforth SLM) [2]. As a model of the interaction of two languages in one subject, learned sequentially, the SLM is potentially generalisable to simultaneous bilinguals; this extension is the hypothesis tested here. The accuracy of the SLM's predictions is tested against data from a study of bilingual (French-English) children.

Three of the main tenets of the SLM are that (i) bilinguals cannot fully separate their two languages' phonetic subsystems; (ii) discrepancies between L1 and L2 learner's productions are due to cross-language interference either through assimilation or dissimilation, (iii) mastery of L2 production depends on prior mastery of perception. Recent research by Flege and colleagues has also suggested an important role for L2 learners for the degree of usage of L1 [8]; for simultaneous bilinguals relative usage may also affect the degree of interference in each language. Recent studies [1,6] suggest that (i) is borne out for simultaneous bilinguals; they have been found to develop phonetic categories in their two languages which although distinct from one another, are less so than is the case for matched monolingual groups. This study tests this further, and examines (ii) by comparing interference in production and perception in the same group and (iii) by examining groups living in France and England, hence exposed differentially to the two languages.

The focus of the study is the voicing contrast. VOT has often been measured in L2 studies. As French has a lead/short-lag contrast whereas English contrasts short- and long-lags, VOT is also an appropriate measure here. Alongside it, the association between voicing category and f0 onset frequency is measured [3,4] (henceforth 'f_0 parameter'); this parameter differs cross-linguistically, being important in French, but absent or intermittently used in British English [7]. For both these parameters, clear developmental trends have been demonstrated in monolinguals; VOTs become shorter throughout childhood, while the f_0 parameter is learnt between 6 and 11.

The subjects investigated are monolinguals and simultaneous bilinguals aged 6 and 10. Within the bilingual subjects are distinct groups living in France and England to allow testing of the hypothesis that interference is modulated by relative degree of language usage.

Applying the SLM, we hypothesize that simultaneous bilinguals will, for VOT (i) be unable to maintain complete separation of their languages; (ii) show effects of assimilation between the two; (iii) show assimilatory effects to different degrees according to degree of exposure to their languages. However, as the f_0 parameter is only used in French, we hypothesize that bilinguals will NOT transfer it to English.

2. SUBJECTS

Child subjects were recruited through schools. In the case of monolinguals, these were primary schools in Southern Britain and the Paris region. Bilinguals were studying at schools either in London (below and in figures referred to as the 'ER' group), or Paris (henceforth the 'FR' group). All bilingual students had been exposed to both languages from birth in the home and were receiving some formal instruction in each. They were selected by teachers as being those who were held to be performing to normal monolingual standards in both languages. In addition to the child subjects, to allow the establishment of adult norms for each language, groups of monolingual adults participated in both experiments. With the exception of one adult English monolingual, all subjects participated in both experiments. Table 1 shows the number of subjects in each group.

Table I: Number of subjects per group

	English Mono	ER	FR	French Mono
Adult	5			5
10 years	6	9	10	7
6 years	5	5	6	8

3. EXPERIMENT 1

3.1. Method

Utterances of the words '*cash*' and '*gash*' in English and '*cache*' and '*gâche*' in French were recorded, these being elicited using flash cards with appropriate pictures. Monolinguals produced only the words in their mother tongue, bilinguals produced both sets on different days, following an initial session to establish the correct language mode. Subjects were first familiarized with the pictures and prompted to produce the target words without actually being told what they were. If a child seemed not to know the word, it was taught. A number of other words were also elicited, as distractors. When all words were clearly known, the experiment proper began. Each word was elicited six times in total, the pictures for all words being presented in random order.

The recordings were then digitized and analysed using PRAAT. For each subject, the first five satisfactory utterances of each target word were measured. VOT was measured from the release burst to the onset of energy in the first formant, which gave the most consistent measure for the child subjects, who often produced unclear, breathy onsets. f_0 onset frequency was measured in the first pitch period following voicing onset. Ohde [4,5] has shown that this pitch period contains the strongest effects associated with voicing category.

Means for each subject were taken across the five utterances for VOT and f_0 onset frequency for both voiced and voiceless tokens. For voiced tokens, this means that positive and negative VOTs were sometimes averaged together; see Watson (submitted) [9] for an analysis of the problems with this procedure and a proposal to modify it. Repeated measures analyses were then carried out for (i) the monolinguals across languages; (ii) all speakers of each language separately; (iii) the bilinguals across their languages (these analyses cannot be combined, as language is a between-subjects factor for monolinguals, but within-subjects for bilinguals). As the f_0 and VOT measures are potentially correlated, an initial multivariate analysis was carried out in each case, followed by univariate analyses of each measure. The factors examined were (i) voicing category; (ii) age; (iii) bilingualism and (iv), for bilinguals, country of residence.

3.2. Results

The VOT results (Figure 1) confirm the predictions made in section 1. The bilinguals make significant distinctions between the two voicing types ($F[1,33]=295.8$, $p < .001$) and differ between English and French (($F[1,33]=175.1$, $p < .001$). They differ, however, from the monolinguals in both languages, $F[1,48] = 12.69$, $p < .00$ for English and $F[2,42] = 5.702$, $p < .01$ for French. The bilinguals' VOT values for the two languages are closer together than the monolinguals', especially for voiced tokens (predictions (i) and (ii)). There is a significant difference between the ER and FR groups ($F[1,33] = 5.37$, $p < .05$); overall the ER group is thus somewhat more monolingual-like than the FR group in English and vice-versa (prediction (iii)).

Comparisons across age groups strengthen the evidence of an effect of country of residence; the effect appears to grow stronger with age. The French /k/ productions of the 10 year old ER group were significantly longer than those of the monolinguals. This is neither true of the 10 year old FR group, nor of either of the 6-year old groups. The same group contrasted with all other bilingual groups in having predominantly short-lag /g/ tokens in English, not statistically distinguishable from those of monolinguals. The results suggest an age-related diminution in interference in English for the ER groups but an increase in the interference they show in French.

The third graph (Figure 2) shows, for each group and language, the difference in f_0 onset frequency between /k/ and /g/ for the first pitch period. The prediction for this parameter was that there would be no interference effects; this holds only for the ER bilingual groups: they follow the monolinguals in showing no voicing contrast related differential in English, but showing one in French from the age of 10. The 10 year old FR group, however, produce significant differences in both languages (t[9] = 3.24, $p < .01$ in English, t[9] = 2.72, $p < .05$ in French). f_0 production data thus confirms that bilinguals cannot reliably avoid cross-linguistic interference, but that the extent of this is dependent on degree of exposure to each language.

4. EXPERIMENT 2: PERCEPTION

4.1. Method

Subjects performed identification tests on two VOT continua, one (the basic continuum) with an f_0 onset of 128 Hz, the other (the low f_0 condition) with an f_0 onset of 103 Hz. The continua were produced using the Klatt synthesizer, and represented the same words as used in Experiment 1, viz. '*cash/gash*' and '*cache/gâche*'. For each continuum, two different language conditions were produced by appending English and French carrier phrases. Monolingual subjects then performed the tests in their one language, bilinguals in both languages on different days, preceded by a discussion in the appropriate language. Labelling curves were then obtained for each subject's responses to each continuum using logit analysis, and the parameters of 50% crossover point and slope extracted. To determine if significant cue-trading had occurred, repeated measures analyses were performed comparing the basic and low f_0 conditions with the same factors as for experiment 1. Crossover and slope differences across the two continua were then calculated for each subject and these were input to multivariate analyses, to allow monolingual-bilingual comparisons.

4.2. Results

The slope of the identification curves proved to have little statistically significant association with bilingualism, so only the results for crossover are given below. For the basic continuum, these are shown in Figure 3. Predictions (i) – (iii) are again confirmed. Although bilinguals have distinct crossover values for their two languages, those values diverge less than those of monolinguals; again, this divergence is only significant, for the language of the country in which the bilinguals are *not* resident. The FR group differs from the English monolinguals in both age groups (F[1,11] = 6.14, $p < .05$ for the 6 year olds, F[1,15] = 9.27, $p < .01$ for the 10 year olds), while the ER group differ significantly from the French monolinguals only at age 10 (F[1,15] = 5.2, $p < .05$). The latter result is caused by changes in variability rather than a clear progression in interference.

Figure 1: Average VOT values for /k/ (upper) and /g/ (lower)

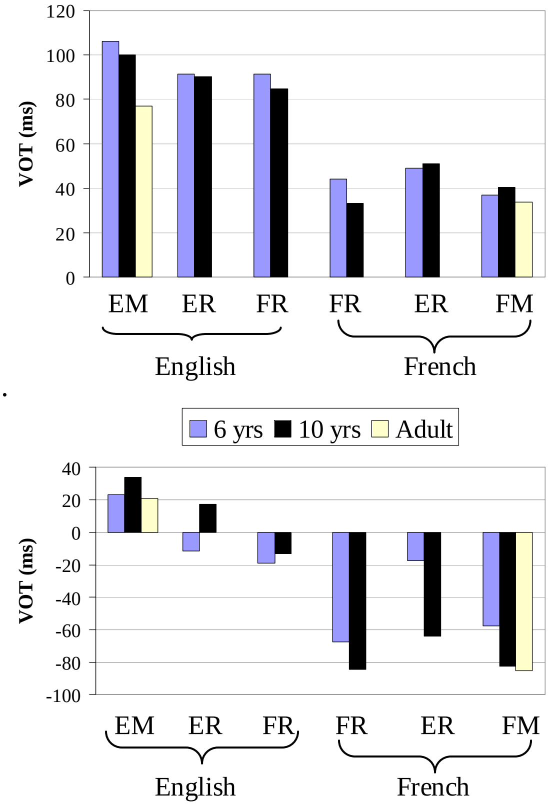
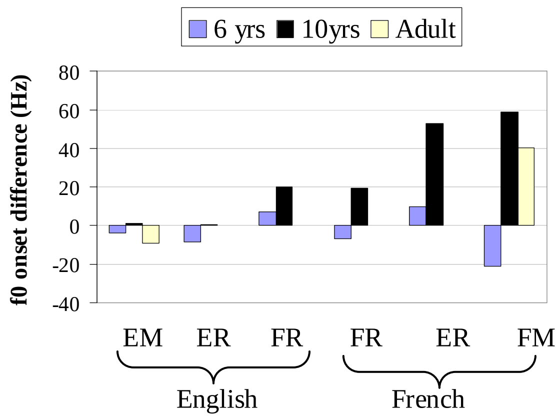


Figure 2: Average /k/-/g/ F0 onset differences.



The VOT/ f_0 cue-trading effect is shown in Figure 4. Here also, effects of bilingualism interact with country of residence. For the monolinguals, the trading effect was non-significant for all English groups, but significant for French adults and 10 year olds. The bilinguals showed interference in both languages. In English, all except the 10 year old ER group showed a significant trading effect (unlike the monolinguals); furthermore all, with the exception of the same group, differed significantly from the age-matched monolinguals ($F[1,41] = 24.176$, $p < .001$). In French, the six year olds overall and both ten year old groups showed an effect, while only the 10 year old ER group differed significantly from the monolinguals ($F[1,15] = 5.03$, $p < .05$), the former showing a smaller, though significant trading effect. There is thus evidence that the bilinguals extend the use of this feature from French to English, but that the older English monolinguals abandon it, with a concomitant weakening of their use of it in French.

Figure 3: Average Identification curve crossover

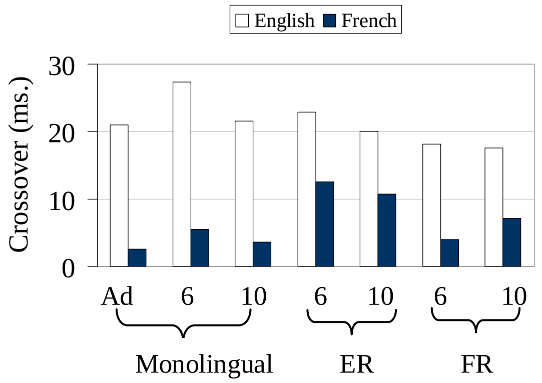
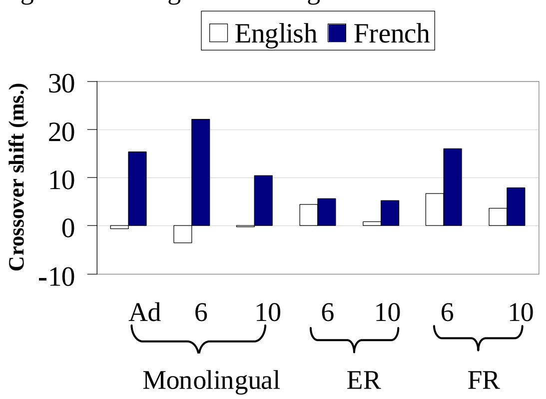


Figure 4: Average cue-trading shift in crossover



5. DISCUSSION

In both experiments the three predictions arising from the SLM were borne out. Bilinguals showed interference in both production and perception, due to cross-linguistic interference. The extent of this interference depended on their relative exposure to their two languages. Interference could diminish as well as increase with age. In Experiment 1 the 10 year old ER group had moved towards the use of short-lag voiced stops in English. In Experiment 2, the same group had learnt *not* to use the f_0 parameter in English. Comparison between the experiments also confirms that perception may lead production in the acquisition of certain parameters, and their transfer across languages. 6 year old subjects showed no consistent use of the f_0 parameter in production, but in perception such use was found in both languages. In the use of this feature, the bilinguals are in advance of the monolinguals; the monolingual French group showed no such effect at this age. A hypothesis to be explored in future work is that exposure to two languages makes children more attentive than monolinguals to voicing features other than VOT.

6. CONCLUSION

The two experiments reported here tested the applicability of predictions derived from the SLM to simultaneous bilinguals. The predictions were supported, although these subjects show varying patterns of precocious interference that may not be found in L2 learners. The importance for simultaneous bilinguals of relative exposure to each language is confirmed. We conclude that this factor should be incorporated into an adapted version of the SLM to form the basis of a general theory of bilingual phonological acquisition.

7. REFERENCES

[1] Bosch, L and Sebastien-Gallés, N. 2003. Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and Speech,* 46 (2-3), 217-243.

[2] Flege, J.E. 1995. Second language speech learning. In Strange, W. (Ed.) *Speech perception and linguistic experience.* Timonium MD: York Press

[3] Haggard, M., Summerfield, Q., & Roberts, M. 1981. Psychoacoustical and cultural determinants of phoneme boundaries: evidence from trading F0 cues in the voiced- voiceless distinction. *JPhon* 9, 49-62.

[4] Ohde, R.N. 1984. Fundamental frequency as an acoustic correlate of stop consonant voicing. *JASA* 75, 224-230

[5] Ohde, R.N. 1985. Fundamental frequency correlates of stop consonant voicing and vowel quality in the speech of preadoloescent children. *JASA* 78, 1554-1561.

[6] Sundara, M., L. Polka and S. Baum 2006. Production of coronal stops by simultaneous bilingual adults. *Bilingualism: Language and Cognition* 9 (1) 97-114.

[7] Watson, I.M.C. (1987) Children, F0 and the voicing contrast; a cross-linguistic study. *Progress Reports from Oxford Phonetics* 2, 137-148.

[8] Flege, J., Frieda, E., & Nozawa, T. 1997. Amount of native language (L1) use affects the pronunciation of an L2. *JPhon* 25, 160-186.

[9] Watson, I.M.C. Acquisition of the voicing contrast by simultaneous bilinguals. Submitted to *Journal of Phonetics.*