The Influence of Language Dominance on Bilingual VOT Production: A Case Study*

Laura E. Herbst

Max Planck Institute for Psycholinguistics, Nijmegen

Longitudinally collected VOT data from an early English-Italian bilingual who became increasingly English-dominant was analyzed. Stops in English were always produced with significantly longer VOT than in Italian. However, the speaker did not show any significant change in the VOT production in either language over time, despite the clear dominance of English in his every day language use later in his life. The results indicate that – unlike L2 learners – early bilinguals may remain unaffected by language use with respect to phonetic realization.

1 INTRODUCTION

In second language (L2) acquisition research, relative frequency of native language (L1) and L2 use have been found to have an impact on phonetic realization, in both L1 and L2. For example, Piske, MacKay and Flege (2001; see also Flege, MacKay & Piske 2002) compared the effect of self-reported L1-use on foreign accent in two groups of native Italian speakers of English, living in Canada: a group of early L2-learners, who had begun learning English as children around the age of 7 or 8, and a group of late L2 learners who had begun learning it only as adolescents or adults. The researchers found that L1-use had an independent effect on accent in the L2 in both groups alike. Similar results have been reported by Flege, Yeni-Komshian and Liu (1999), and Guion, Flege and Loftin (2000).

Piske et al. (2001) suggest that effects of language use are a result of the phonetic input on the structure of the "composite (phonetic) representations", which L2 learners are said to develop according to the Speech Learning Model (Flege, 1995). In this model, it is assumed that L2 learners are prevented from forming a new phonetic category for a given L2 sound if they perceive this sound as being similar to a sound in their L1. Instead, 'a category that subsumes the equated L1 and L2 speech sound will develop over time' (MacKay, Flege, Piske, & Schirru, 2001: 516), which reflects the phonetic properties of both. As a result, the production (and perception) of both sounds will be a "compromise" between the two. This single representation would then get skewed towards the norm of the language used more frequently.

However, there are findings which show that some L2 learners are capable of forming a new phonetic category, but nevertheless show language use effects: Sancier and Fowler (1997) found that even a temporary change in the language environment of an L1-Portuguese learner of English influenced stop consonant production: This speaker's Voice Onset Time (VOT) in the realization of VOICELESS[1] bilabial and alveolar stops in both languages were longer after a 4-month stay in the US, and shorter after a 2.5-month stay in Brazil. The effects were also correlated with the perception of an accent in Portuguese by native speakers. Despite having learnt English at the age of 15, the speaker differentiated VOT values for both languages, but both were equally

* I would like to thank Bob Ladd for his expertise and great support, and for kindly providing me with the data, without which this project would not have been possible.

[1] Upper case is used to indicate the phonemic status of a stop.

CamLing 2006: 100-107.

affected by the speaker's linguistic environment. In addition to having shown that language-use effects may also occur in speakers who have command over two phonetic representations, Sancier and Fowler's study also established that VOT is a phonetic parameter that is affected by relative frequency of language use (and which contributes to the perception of accent). Furthermore, while language-use effects have mostly been discerned only by making observations across populations, Sancier and Fowler showed that such changes can occur within a single speaker as well.

Against this backdrop, it is quite remarkable that there is up to this point no study that has looked into the possible effects of language use on phonetic realization in *first-language* bilingual speakers, that is, speakers who have been exposed to two (or more) languages from birth. Studies on first-language (or "early") bilingual phonology are mostly centred on the actual acquisition process of two distinct systems in very young children, the main finding being that language differentiation usually takes place very early (e.g., Deuchar & Quay, 2000; Johnson & Wilson, 2002). The stability of the acquired phonologies in view of changing linguistic environments has not received much attention.

The present study aims at shedding some light on this question by following the VOT development of an early English-Italian bilingual over the period of 6 years (age 8 to age 14). With the exception of Sancier and Fowler's (1997) study, in which the subject was tested twice, most studies on language use have worked cross-sectionally. In contrast to this, the present study is based on truly longitudinal data, which can be matched against changes in the child's language environment and use.

2 METHOD

2.1 Subject

The subject of this study, M., is the son of a native (American) English-speaking father and a native Italian-speaking mother, and has a brother four years younger than him. Both parents speak the other language as a second language, but the mother uses English much more extensively than the father uses Italian. The family live in Edinburgh (Great Britain), and the parents usually speak English with each other.

The child has had bilingual input from birth. However, his first years of life may be described as more Italian-dominant. Before M. started school at the age of 5;0, he spent half of the time in an English-speaking nursery, and the other half with an Italian child minder. He continued interacting with the Italian child minder also when he entered school (English-speaking), and used Italian with his mother and his younger brother. Judging from the parental estimate of use of English and Italian at that time, one could assume that M. should have been relatively balanced.

He started going to secondary school (also English-speaking) at age 11;00, and it is around this age that it may be adequate to describe M.'s language use/environment as becoming increasingly English-dominant. Some time between the ages of 11;0 and 12;6, M. gradually began to shift to speaking exclusively English with his brother. However, even though M., who is now 15, may be rather English-dominant at this point, he still uses Italian every day with his mother.

2.2 Data collection

The data consist of a series of audio recordings made under studio conditions. In each recording session, M. read a couple of texts from children's books, one from each

language.[2] On two occasions, an additional recording was made in which M. read an Italian text with an imitated English accent.

2.3 Acoustic analysis (VOT measurement)

The analysis was restricted to VOICELESS bilabial, alveolar and velar stops. Only singletons in the onset of stressed syllables that occurred in word-initial or intervocalic position were analyzed. VOT was measured from the beginning of the release burst to the point at which the waveform sets off from the baseline to go over into a more sinusoidal shape, coinciding with the onset of the lower formants and/or the onset of clear vertical striations in the spectrogram.

2.4 Other information

Each stop was coded for its place of articulation, POA (bilabial/alveolar/velar), as well as for the type of the following vowel (high/non-high). For both languages, /ɪ,ʊ / and the corresponding approximants /j,w/ were taken to be high, all others as non-high. The stop's position within the word (initial/medial) was also noted. Furthermore, speaking rate (SR) was measured.[3]

2.5 Hypotheses

Italian VOICELESS stops are produced in the short-lag region (i.e., voicing sets in at around 0 to +25 milliseconds (ms) after the opening of the consonantal closure). The corresponding stops in English are typically produced in the long-lag range (i.e., voice onset at around +60 to +100ms). If language use has the same effect on first-language bilinguals as it has been found for L2 learners, one would expect to observe a change in M.'s VOT production as he became increasingly English-dominant. More specifically, one would expect to find an increase in VOT at least in his Italian stops, but possibly also in his English stops. Alternatively, the relative frequency of use of the two languages may not have any influence at all. In this case, M.'s VOT production would be expected to remain constant over time.

3 RESULTS

3.1 English

In the English sample, 790 stops were analyzed. The factors POA (bilabial/alveolar/velar), VOWEL (high/non-high) and POSITION (initial/medial) were entered into an ANOVA, with VOT as the dependent variable (DV). There was a main effect of POA [$F(2, 779) = 4.301$, $p < .05$] and of VOWEL [$F(1, 779) = 8.052$, $p < .05$],

[2] Investigating VOT was not the primary purpose of the data collection, which is why the material was not balanced across languages with respect to quantity, and why it had not been used consistently.

[3] A measure for SR was obtained by dividing the number of syllables in the text by the overall duration of the recording. High SR does not necessarily entail a high articulation rate (AR; cf. Laver 1994: 539ff). However, track was also kept of the number of incomplete closures. These were significantly correlated with the SR measure, pointing at a link between SR and AR.

but also a significant interaction between these two [$F(2, 779) = 3.470$, $p < .05$]. POSITION had no effect.

Post-hoc *t*-tests revealed that overall, mean VOT in bilabials ($M = 43.65$ms) was significantly shorter than in alveolars ($M = 60.66$ms; $p < .001$), and significantly shorter than in velars ($M = 62.07$ms; $p < .001$). Mean VOT with high vowels ($M = 72.31$ms) was significantly shorter than with non-high vowels ($M = 52.37$ms; $p < .001$). The interaction stems from the fact that the POA effect depends on the type of vowel a stop is combined with. In contrast to this, vowel type effects occurred independently of POA. VOT in stops preceding high vowels was always longer than in the non-high equivalents.

VOT turned out to be very weakly but significantly correlated with speaking rate [Pearson's $r = -.071$, N = 790, $p < .05$], which means that VOT decreased as speaking rate increased.

The analyses have established that M.'s English VOT production was affected by intrinsic and contextual factors. This means that any developments of mean VOT that may be observed across time may merely be a result of different compositions of the samples with respect to the distribution of the various stop + vowel combinations. It is not possible to simply compare the overall means of all dates with each other when looking for temporal developments. Instead, these effects have to be controlled for somehow in order to make VOT comparable.

The data was therefore aggregated in the following way: For each date, the mean for each of the six stop + vowel combinations (henceforth "phone") has been calculated. The obtained values were treated as single data points. Mean VOT for each date was then computed from these six values. This way each phone contributes equally to the overall mean of a given date, and none of them can exert an undue influence, yet if there were genuine changes in M.'s VOT production, these should be reflected in significant differences between these aggregated means.

A one-way repeated measures ANOVA was then carried out with mean VOT as DV and DATE as within-subjects factor. DATE did not produce a main effect [$F(2.716, 13.579) = 1.968$, $p = .170$]. In other words, M.'s phonetic realization of his English stops did not change over the period covered by the recordings.

However, it is conceivable that there has been some development with respect to the overall variability in M.'s production. A measure for variability was obtained by dividing the standard deviation for each of the six phones at a given date by the respective mean, which gave a proportional value of variation for this phone. The mean percentage of variation around the mean for each date was then calculated by averaging across these six values.[4] The overall percentage of variation around the mean was 37.66%. The correlation between DATE and VARIATION was not significant [Spearman's $\rho = -.440$, N = 13, $p = .133$], which means that the variability in M.'s English stop production remained constant.

3.2 Italian

For Italian, 1,206 stops were analyzed. As with English, the factors POA, VOWEL and POSITION were entered into an ANOVA with VOT as DV. There were main effects of POA [$F(2, 1194) = 69.769$, $p < .001$] and of VOWEL [$F(1, 1194) = 105.551$, $p < .001$],

[4] The standard deviation was not used, because it always has to be looked at relative to the mean. Since Italian VOT values are likely to occupy a different range on the VOT continuum, the standard deviations in English would be larger than in Italian in absolute values. This would in turn make a comparison across languages difficult.

as well as an interaction of these two factors [$F(2, 1194) = 6.167$, $p < .05$]. There was no effect of POSITION.

Post-hoc tests showed that alveolars had the shortest VOT (M = 30.97ms), followed by bilabials (M = 33.79ms; difference significant at $p < .05$), which had in turn significantly shorter VOT than velars (M = 48.45ms; $p < .001$). Mean VOT was significantly longer for stops occurring with high vowels (M = 47.61ms) than for those that were followed by non-high vowels (M = 30.59ms; $p < .001$). This vowel effect was independent of the POA. The interaction is due to the fact that the POA effect was not the same for high and non-high vowels.

A correlation run between VOT and speaking rate did not produce a significant result [Pearson's r = .028, N = 1,206, p = .326].

Following the data aggregation procedure described for English above, the normalized data were subjected to a one-way repeated measures ANOVA with mean VOT as DV, and DATE as within-subjects factor. An effect of DATE was found: $F(3.404, 17.021) = 3.170$, $p < .05$. Post-hoc Bonferroni t-tests showed that this was due to the fact that mean VOT was significantly shorter on the last recording on 28/03/2005 (33.90ms) compared to mean VOT on the 20/08/2004 (41.15ms) at $p < .001$. However, as this was the only significant difference in the entire data set, and as these two dates were not even subsequent, this does not seem to constitute any clearly discernible general trend over time.

Variability of production was examined the same way as for English. The variability in the entire data set was 37.86%, a value comparable to that in English. Unlike in English, however, variability decreased in Italian over time, as evidenced by the significant negative correlation between DATE and VARIATION [Spearman's ρ = -.634, N = 14, $p < .05$].

3.3 "Accented Italian"

The same analyses as with English and Italian were run over the "accented Italian" data (62 stops analyzed). The only effect found was one of VOWEL, [$F(1, 54) = 10.050$, $p < .05$]. A post-hoc comparison showed that VOT was significantly longer with high vowels (M = 84.41ms) than with non-high ones (M = 61.90ms; $p < .001$).

3.4 Cross-language comparison

So far, it has been shown that in both languages VOT was affected by vocalic environment and POA. It has further been shown that in neither language was there any indication for genuine changes in VOT production over time. The question that has not yet been answered is to what extent English and Italian VOT differ from each other.

Since no temporal developments were found, the time factor was ignored in the subsequent cross-language comparison. VOT in both languages was compared using the same "normalization" procedure as before. The average VOT for the (normalized) English sample was 60.95ms (SD = 14.46ms), for Italian it was 39.03ms (SD = 11.85ms). This difference was statistically significant [$t(10) = 2.872$, $p < .05$].

Mean VOT in the "accented Italian" after normalization was 71.56ms (SD = 16.00ms), which was significantly longer than VOT in the Italian data ($t(10) = 4.001$, $p < .05$), but did not differ significantly from English VOT [$t(10) = 1.205$, p = .256]. When speaking Italian with an English accent, M. produced VOT values that were clearly closer to his English stops than they were to his Italian stops.

4 DISCUSSION

The central question of this study was whether changes in language dominance can induce changes in the VOT production of a first-language bilingual child. First of all, it can be noted that M.'s VOT production exhibits all the characteristics discussed in the literature. There were strong effects of vowel context and POA in both English and Italian. Subsequent high vowels always led to longer VOT than non-high ones, which is in accordance with a number of other studies (e.g., Esposito, 2002; Klatt, 1975). Velar POA yielded the longest VOT values, as has been reported before (cf. Cho & Ladefoged, 1999). Also in line with previous findings (e.g., Kessinger & Blumenstein, 1997; Magloire & Green, 1999), long-lag stops were produced with slightly longer VOT at higher speaking rates. The results show that M.'s production patterns are absolutely "normal". Considering the striking regularity of the data, there is nothing to suggest that what has been found is "atypical" in some way.

Regarding the key issue, namely the question of whether M. shows any signs of changes in his VOT patterns caused by increased use of English, the statistical analyses have not brought about evidence for any genuine changes in either language. M. has had command over fully differentiated phonetic realizations[5] for each phoneme from the first recording onwards and there is no indication that these have undergone any alterations over time. What is more, M. does not only differentiate English and Italian phonemes, he also seems to be aware of the phonetic differences, as reflected in his English-like realization of Italian stops.

Although M. has been using English much more frequently for the past two to three years, and although he also seems to prefer to speak English, this has not caused any restructuring on the phonetic level, neither in Italian nor in English. The increasing use of English has not had a "destabilizing" effect on M.'s production, either; the variability in English remained constant, while it even decreased in Italian.

The results contrast with those obtained for second language learners, who have been found to show "drifts" in their phonetic realizations towards that of the language used more frequently (Flege et al., 2002; Flege et al., 1999; Guion et al., 2000; Piske et al., 2001; Sancier & Fowler, 1997). The present study clearly shows that such effects are not inevitable for first-language bilingual speakers. M.'s case has demonstrated that the full retention of two voicing systems is possible also in view of changes in first-language bilinguals, but it adds to the large body of evidence that learning a language "on top" of another one is not tantamount to learning two languages simultaneously from birth, especially for the domain of phonetics and phonology.

Numerous studies have shown that the foundations for these aspects of language competence in particular are laid very early in life. For example, infants as young as two months are able to recognize global prosodic features of their native language (Mehler et al., 1988), and by the age of nine months they show clear preferences for their L1's phonotactics (Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993). The ability to discriminate non-native phoneme contrasts ceases around ten months (e.g., Polka & Werker, 1994; Werker & Tees, 1984). Phonetics and phonology are likewise those domains for which most researchers agree on a critical or at least sensitive period prior to which a learner must be exposed to a language in order to attain native-like competence. In a review of research on maturational constraints in language learning, Long (1990) comes to the conclusion that in order to have a chance of avoiding a foreign accent, the onset of immersion has to take place before the age of six.

[5] It is, however, possible that his English values are somewhat shorter than those of monolingual English speakers, and his Italian ones somewhat longer than those of monolingual Italian speakers, but this is not of interest for this study.

The results obtained in this study suggest that age may not only be an important factor for the ability to develop target-like (distinct) phonetics in the first place, but that it may also be crucial for the stability of these representations in view of changes in the linguistic environment. Several studies have found some non-first-language bilingual speakers to also differentiate between L1 and L2 voicing categories (e.g., Sancier & Fowler, 1997; Flege 1991) but no longitudinal study has yet presented evidence that these remain in a steady state as language use patterns change. On the contrary, Sancier and Fowler's (1997) result showed that such speakers are apparently susceptible to those changes.

M., on the other hand, who has had input from both languages from birth, has maintained the same VOT patterns despite the disparity in the use of the two languages. Corroborative evidence that age may play a key role regarding the stability of phonetic and phonological competence comes from studies on the benefits of overhearing a language in childhood; that is, from individuals who have been exposed to a language in infancy or early childhood, but who have not acquired it and who have by and large been detached from that language thereafter. It seems that childhood overhearers can maintain native-like competence not only in the perception (e.g., Tees & Werker, 1984; Oh, Jun, Knightly & Au, 2003), but even in the production of phonetic contrasts of the language they have had no contact with for many years (Knightly, Jun, Oh & Au, 2003; see also Au, Knightly, Jun & Oh, 2002).

These findings suggest that even a drastic change in the linguistic environment may not alter or erase phonological or phonetic "traces" if these have been "imprinted" in infancy and early childhood, even if the respective language has not been fully acquired, and could explain the absence of any effect of language use in M.'s VOT production of VOICELESS stops.

This has been the first study to broaden the scope of research on language-use effects on the phonetic level to first-language bilinguals. However, a number of open questions remain. Reporting only on a single case, this study could only show that a shift in language dominance did not lead to VOT modifications in this particular child under these particular conditions. Nothing can be said about other first-language bilingual speakers, or about what developments may have occurred, had M. had a different language history.

For example, the timing of changes in the linguistic environment may play a role. M. was on the verge of puberty when he became more English-dominant. Maybe his production patterns would have developed in a different way if this shift had taken place earlier. The findings from child overhearers suggest otherwise, but it should also be borne in mind that they still constitute a different population from genuine first-language bilinguals.

Additionally, M. has been English-dominant for three years, and is still speaking Italian on a daily basis. It has yet to be investigated what the consequences are of a permanent dominance in one language of maybe ten or even twenty years for these speakers, and what happens if a (fully competent) speaker is completely severed from one language over a long period of time.

Furthermore, there is the possibility that whether or not voicing systems are prone to influences from language dominance depends partly on the proximity of the voicing categories in the two languages. The phonemic categories occupy relatively different ranges of the VOT continuum in English and Italian – one might ponder whether the same results would be obtained for German-English bilinguals, for instance. Future research will have to show to what extent the findings reported here can be generalized to other first-language bilinguals with different language biographies.

REFERENCES

Au, T. K. F., Knightly, L. M., Jun, S. A., & Oh, J. S. (2002). Overhearing a language during childhood. *Psychological Science, 13*(3), 238-243.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics, 27*(2), 207-229.

Deuchar, M., & Quay, S. (2000). *Bilingual acquisition: Theoretical implications of a case study*. Oxford: Oxford University Press.

Esposito, A. (2002). On vowel height and consonantal voicing effects: Data from Italian. P*honetica, 59*(4), 197-231.

Flege, J. E. (1991). Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a 2nd language. *Journal of the Acoustical Society of America, 89*(1), 395-411.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience* (pp. 233-273). Baltimore: York Press.

Flege, J. E., MacKay, I. R. A., & Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics, 23*(4), 567-598.

Flege, J. E., Yeni-Komshian, G. H., & Liu, S. (1999). Age constraints on second-language acquisition. *Journal of Memory and Language, 41*(1), 78-104.

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

Johnson, C. E., & Wilson, I. L. (2002). Phonetic evidence for early language differentiation: Research issues and some preliminary data. *The International Journal of Bilingualism, 6*(3), 271-289.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M. I., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language, 32*(3), 402-420.

Kessinger, R. H., & Blumstein, S. E. (1997). Effects of speech rate on voice-onset time in Thai, French, and English. *Journal of Phonetics, 25*(2), 143-168.

Klatt, D. H. (1975). Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research, 18*(4), 686-706.

Knightly, L. M., Jun, S. A., Oh, J. S., & Au, T. K. F. (2003). Production benefits of childhood overhearing. *Journal of the Acoustical Society of America, 114*(1), 465-474.

Laver, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press.

Long, M. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition, 12*, 251-285.

MacKay, I. R. A., Flege, J. E., Piske, T., & Schirru, C. (2001). Category restructuring during second-language speech acquisition. *Journal of the Acoustical Society of America, 110*(1), 516-528.

Magloire, J., & Green, K. P. (1999). A cross-language comparison of speaking rate effects on the production of voice onset time in English and Spanish. *Phonetica, 56*(3-4), 158-185.

Mehler, J., Jusczyk, P., Lambertz, G., Halsted, N., Bertoncini, J., & Amieltison, C. (1988). A precursor of language acquisition in young infants. *Cognition, 29*(2), 143-178.

Oh, J. S., Jun, S. A., Knightly, L. M., & Au, T. K. F. (2003). Holding on to childhood language memory. *Cognition, 86*(3), B53-B64.

Piske, T., MacKay, I. R. A., & Flege, J. E. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics, 29*(2), 191-215.

Polka, L., & Werker, J. F. (1994). Developmental changes in perception of non-native vowel contrasts. *Journal of Experimental Psychology - Human Perception and Performance, 20*(2), 421-435.

Sancier, M. L., & Fowler, C. A. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics, 25*(4), 421-436.

Tees, R. C., & Werker, J. F. (1984). Perceptual flexibility - Maintenance or recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology - Revue Canadienne de Psychologie, 38*(4), 579-590.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception - Evidence for perceptual reorganization during the 1st year of life. *Infant Behavior & Development, 7*(1), 49-63.

Laura E. Herbst

Max Planck Institute for Psycholinguistics
P.O. Box 310
6500 AH Nijmegen, The Netherlands

laura.herbst@mpi.nl