ACQUISITION OF VOT, AND VOWEL LENGTH BY ENGLISH-GERMAN BILINGUALS: A PILOT STUDY

Nicole Whitworth

Abstract

This paper reports the findings of a pilot study investigating bilingual acquisition of segmental temporal patterns. The productions of two German-English bilingual children, aged 9;11 and 12;5, have been analysed with regard to voice onset time (VOT) and extrinsic vowel length. There are three conclusions to be drawn from this investigation: Firstly, there is a great deal of variation between the subjects. Secondly, bilinguals cannot be treated as two monolinguals in one person. Thirdly, the results indicate that bilingual children are aware of fine-grained variability in the linguistic input they receive, and are able to produce corresponding patterns which are, however, not identical to the adult monolingual target. This provides further evidence in favour of the dual systems hypothesis. It is suggested that the 1/2 systems hypothesis as applied in the field of bilingual language acquisition might be misleading and monolingual and bilingual first language acquisition might result in similar cerebral representation, since within-language variation is likely to be represented by the same neurological structures as between-language variation.

1. Introduction

The results reported in this paper are part of an ongoing project investigating the acquisition of speech timing by German-English bilingual children aged 2;0, 5;0, and 11;0. The variables targeted in the project are initial voice onset time, extrinsic and intrinsic vowel duration, and speech rhythm.

Its aim is to contribute to the ongoing one-or-two-systems debate of bilingual language acquisition by integrating findings and methods from acoustic phonetics, phonology, sociolinguistics, and bilingual and monolingual first language acquisition. The originality of the investigation lies firstly in its general approach, as acoustic studies of bilingual speech on larger samples have rarely been performed in order to address the one/two systems issue. Secondly, in contrast to previous studies this study investigates bilingualism of two closely related languages.

The close linguistic relationship of German and English is the crucial factor utilised to look into the dual/unitary system issue. If it could be shown that the bilingual children systematically distinguish between fine-grained language-specific temporal patterns, this would suggest the existence of two systems rather than a single one, since one might hypothesis that low-level allophonic differences between a bilingual's two languages would be merged if they were represented by a single system.

2. Background

2.1. Bilingual acquisition of speech timing

Children who acquire two languages from birth have to cope with a magnitude of competing, sometimes even contradictory, phonetic and phonological input. To successfully master both languages they have, firstly, to be aware that the auditory signs belong to two different systems and, secondly, to allocate each sign to the appropriate system. In addition to the differences between the two language systems they also have to deal with regional, social, stylistic, pragmatic, and allophonic variants in each language, as well as with intraspeaker variation (i.e. the speech of the same speaker varies with regard to pitch, speaking rate, etc.). This means that on the one hand bilinguals have to associate differences in pronunciation with two underlying unconnected systems, while at the same time taking account of variations in pronunciation which occur within a single system. In spite of this apparently immense challenge, bilingual speakers, who acquire both languages from an early age, seem to achieve monolingual-like pronunciation in both their languages, even to phonetically trained ears.

To be able to speak a language with a non-native accent, a speaker has to have mastered, amongst other things, the correct timing of articulatory sequences. Speech is a temporal event; speech production and perception are both subject to temporal ordering. Speech segments take up a specific period of time, the length or duration[1] of which is conditioned by

- *intrinsic factors*, i.e. segment duration is affected by the articulation of the segment itself (e.g. tense vowels tend to be longer than lax vowels), and
- *extrinsic factors*, i.e. preceding or following segment articulations affect the duration of a segment (e.g. vowels tend to be longer before voiced consonants than before voiceless consonants).

These factors are generally based on physiological and/or physical conditions inherent in the speech production process. However, many factors are also language-specific (Lehiste 1970, Laver 1994, Docherty 1992, Fox 2000) and, thus, the temporal structure of a language has to be learned. This means, that in order to learn a language it is not enough to acquire sets of phonemes and allophones, since it is also the case that fine-grained phonetic detail varies from language to language.

So far research into bilingual acquisition of speech timing has concentrated on bilingual acquisition of voice onset time (VOT). These studies in turn have focused on languages which possess quite different VOT systems, i.e. voicing lead and short-lag VOT vs. short-lag and long-lag VOT, for example Arabic-English (Khattab 1998), Spanish-English (Deuchar & Clark 1996), and French-English (Caramazza et al. 1973, Hazan & Boulakia 1993). In all the studies quoted above the bilingual subjects displayed different production strategies for each of their languages. Unfortunately, there do not appear to be any studies of either intrinsic or extrinsic vowel length in bilingual speech.

2.2. One or two systems?

A bilingual child is exposed to two different sets of language-specific temporal patterns, for example, one German and one English pattern, which vary with regard to VOT, and intrinsic and extrinsic vowel duration. In the case of German and English these patterns make use of the same categories, e.g. both languages use short-lag VOT to mark phonologically voiced and long-lag VOT to mark phonologically voiceless stops, but the specific timing characteristics differ systematically (see section 5.1.2).

How then does a child deal with the different temporal patterns it is exposed to? Does he acquire both patterns, disregard one or the other, or abstract the input to a single mixed system?

[1] Duration/length is defined as “the amount of time taken up by a speech event” (Laver 1994: 431). The terms ‘duration’ and ‘length’ are here used synonymously to refer to the phonetic length of a speech event. ‘Quantity’ on the other hand is used to denote phonemic length.

The question whether bilinguals develop two separate systems or initially a single merged system for both their languages is at the centre of a long-standing debate in the field of bilingual language acquisition. Proponents of the unitary system hypothesis, e.g. Volterra & Taeschner (1978), and Vihman (1985), claim that bilingual children initially organise their two languages in a single merged system, whereas the proponents of the dual systems hypothesis, e.g. Genesee (1989), and Deuchar & Clark (1996), assert that they have two separate language systems from the beginning.

Supporters of the unitary systems hypothesis have generally based their arguments for an initial, mixed system on the existence of language mixing in the early speech of their subjects. Leopold (1947) reports the mixing of English and German items in the vocabulary of his bilingual daughter Hildegard. Volterra & Taeschner (1978) describe the acquisition of two German-Italian bilingual girls. They propose three stages of representation for the girls, with a single lexical system at stage 1, which developed into two systems at stage 2. However, the girls still had a unitary system for syntactic representation at stage two which only diverged later at stage 3. Again evidence of language mixing was used to argue for an initial mixed system.

Genesee (1989) asserted that the dual systems hypothesis cannot be dismissed on the grounds of the existence of language mixing in the children's speech on the basis of a re-examination of the data of the previous studies. He suggested that language mixing may be due to incomplete linguistic repertoires and mixed input. He also showed that the children distinguished between their languages by context to the best of their abilities at each stage of their language development. A study of French-German bilinguals showed that bilinguals consistently apply the correct word order when speaking in one of their languages Meisel (1989).

Schnitzer & Krasinski (1994, 1996) examined the phonological development of their two Spanish-English bilingual sons Fernando and Zevio. An analysis of the speech of their first son Fernando lead them to believe that he initially developed a single mixed system which only later developed into two separate ones. In contrast they found that Zevio developed two phonological systems from the beginning. They concluded that phonological development should be viewed in "terms of acquiring segmental repertoires rather than systems" (Schnitzer & Krasinski 1996: 562).

Finally, Cruz-Ferreira (to appear) questions the validity of the question when bilinguals differentiate their languages. Bilingual children achieve maximum differentiation between their languages through prosodic and phonetic features from an early age within their restricted articulatory abilities.

2. Aims of the Study

If a bilingual realises a different timing pattern for each language, in spite of the minimal differences of the temporal structures of these languages, it would make sense to assume that the two languages are represented by separate linguistic systems. If, however, the languages of a bilingual speaker were represented by a single mixed system, it would be unlikely to find any systematic differences between the utterances of language A and language B. In the latter case the timing pattern might consist of the 'native' system of either language or some kind of mixed system. The study thus aims to answer the following research questions:

- Do bilingual children possess two separate linguistic systems, rather than a single unified system,

- and if so, are these systems identical to those of the respective monolingual speakers?

3. Subjects

Reuben, aged 9;11 at the time of the recordings, and Salome, aged 12;5, are two German-English bilinguals from Bradford, UK. Both children have been exposed to German and English from birth. The parents followed a one-parent-one-language approach; the mother only speaks German to the children and the father only English. The children were both born in the UK, and have been living there ever since. English can be said to be their dominant language since exposure to it has been greatest. English is the main language spoken in the environment, e.g. at school, by friends, etc. German is spoken on the annual holidays (two to four weeks) in Germany, and when the children were younger they regularly (once a month) visited a German-speaking playgroup. When the recordings were made both children had not been to Germany since the previous year.

Table 1 gives a concise categorisation of the children's speech. The judgments of language accent and proficiency have been made by listening to the parents' and children's speech in conversation and are not based on a more detailed phonetic analysis. The parents' linguistic background is summed up briefly in Table 2 below

Table 1: Categorisation of Reuben's and Salome's accents

	Reuben (9;11)	*Salome (12;5)*
German	English accent.	Native-like. Northern Standard German accent.
English	Native-like. Slight Northern/Yorkshire accent.	Native-like. Northern/Yorkshire accent.

Table 2: Parents' linguistic background

	Mother	*Father*
L1:	German - Northern Standard German accent	English. Slight Yorkshire accent
L2:	English. Fluent	German. Intermediate
Other L2	French	French
Occupation	Language teacher (German)	Architect

Both the adults and the children have a very positive attitude towards bilingualism.

4. Data Collection and Material

Tape recordings of the children's speech were made during home visits. The data were elicited by means of a naming task, in which the children were asked to name the objects depicted on flashcards. Recordings were made on a DAT machine with lapel microphones.

Each subject has been recorded on two occasions. In each of these two recording sessions only one language was spoken. The respective monolingual parent was asked to be present, although, this was not always possible. These steps were taken in order to control for language mode, that is the aim was to activate a monolingual German mode in the German recording session, and a monolingual English mode in the English recording session (Grosjean 1998).

The recorded linguistic material consists of one-word utterances with a CVC(VC) structure. Segment durations are measured in the stressed word-initial syllable of the token words, namely

- initial voice onset time (full details in section 5.1),
- intrinsic vowel length (full details in section 5.2), and
- extrinsic vowel length (full details in section 5.3) in each of the two languages.

5. Data Analysis

The recordings were analysed auditorily and then tokens containing the respective target variables were extracted and analysed acoustically. VOT and vowel length were measured on wideband spectrograms produced on a Sensimetrics Speech Station 2.

A number of tokens were rejected for various reasons: voice overlay, nasal release of stops, production of German /r/ as an alveolar approximant, etc. Extreme values have also been discarded in order to render the data more representative, i.e. the two lowest and two highest measurements in each of the four categories were disregarded.

The resulting measurements were ordered into four categories: (1) English voiceless consonants or lax vowels, (2) English voiced consonants or tense vowels, (3) German voiceless consonants or lax vowels, and (4) German voiced consonants or tense vowels. Each of the four categories were compared with each other to see whether the differences between them were significant. The mean duration of each category, and voiceless/voiced and lax/tense ratios was then calculated. Ratios are used to make the adult and child data comparable, since it is well known that segment durations are generally longer in children than in adults (Hawkins: 1979, Stoel-Gammon & Buder: 1999).

5.1. Voice Onset Time (VOT)

5.1.1. Definition

Voice Onset Time (VOT) is defined as the time which elapses from the point of release of a stop closure to the onset of vocal fold vibration. (Lisker & Abramson: 1964). In many languages of the world VOT is used to distinguish between voiced and voiceless sounds. There are three perceptual categories (Laver 1994: 349):

- *voicing lead*, where the onset of phonation occurs before the stop closure is released (by 25ms or more),
- *short-lag VOT*, where the onset of vocal fold vibration occurs simultaneously to the stop closure release (within +/- 20ms), and
- *long-lag VOT*, where the onset of phonation occurs after the stop closure is released (by 25ms or more).

5.1.2. German and English VOT patterns

German and English both use short-lag VOT to mark voiced and long-lag VOT to mark voiceless stops. They differ with regard to the temporal implementation of the voicing categories.

For German voiced VOT Braunschweiler (1997) cites a mean of 16 ms and for voiceless VOT a mean of 51 (when followed by the long or short open central vowel). Docherty (1992) found VOT means of 59 ms for British English voiceless and 32 ms for British English voiced stops. In this study there was also some evidence of free variation between prevoicing and short-lag stops as realisations of phonemically voiced stops. However, for the means reported here voicing lead has been disregarded.

Table 3: Initial VOT in German and English

	English	**German**
Voiceless stops (ms)	59[A]	51 [B]
Voiced stops (ms)	32[A]	16[B]
Ratio voiceless/voiced	1.84[A]	3.19 [B]

[A] Docherty (1992); [B] Braunschweiler (1997)

The ratio of voiceless to voiced VOT in English is 1.84 (Docherty 1992) and 3.19 in German (Braunschweiler 1998). A voiceless/voiced ratio of 1.84 means that in English the delay in the onset of voicing is 1.84 times greater for voiceless stops than for voiced stops. Correspondingly, the German ratio 3.19 means that the VOT in voiceless stops is 3.19 times greater than that for voiced stops.

5.1.3. Analysis

A total of 88 tokens per subject have been analysed with regard to initial voice onset time. VOT was measured from the start of the release burst to the onset of vocal fold vibration indicated by the start of periodic energy in the waveform.

5.1.4. Results

Table 4 and Table 5 show the results for Reuben and Salome respectively. Values for which *t*-tests have shown significant differences between the English and German productions are marked with an asterisk (*). The categories compared are: English voiced VOT vs. German voiced VOT and English voiceless VOT vs. German voiceless VOT.

Reuben and Salome only used short-lag VOT to mark voiced stops and long-lag VOT to mark voiceless stops in both languages. There was no occurrence of voicing lead in the production of the subjects' English productions. Both children differentiate between German and English voiceless ($p \leq 0.01$) stops. Furthermore, Salome realises VOT in English voiced stops differently from German voiced stops ($p \leq 0.02$).

Table 4: Reuben (9;11) initial VOT

	English		**German**	
	voiceless	**voiced**	**voiceless**	**voiced**
No. of tokens	21	33	16	18
Range (ms)	57-88	4-22	66-103	9-24
Standard deviation (ms)	10	5	12	4
Mean duration (ms)	*74.62	13.50	*87.13	14.22
Ratio voiceless/voiced	5.67		6.13	

Table 5: Salome (12;5) initial VOT

	English		German	
	voiceless	**voiced**	**voiceless**	**voiced**
No. of tokens	21	38	13	16
Range (ms)	60-109	4-20	48-75	8-22
Standard deviation (ms)	15	5	9	5
Mean duration (ms)	*84.05	*11.24	*63.31	*14.31
Ratio voiceless/voiced	7.48		4.42	

Figure 1 compares the voiceless/voiced ratios as produced by Reuben and Salome with the ratios found in monolingual adult speech (see Table 3). The first thing to notice is that the ratios of both children are in all cases greater than those of the adults.

Reuben displays a smaller difference between the ratios of the two languages than the monolingual adult target. However, since his English and German mean durations for voicing delay in voiceless stops differ significantly, and that for his voiced stops are similar, it can be concluded that there must be a significant difference between the ratios of both languages. (If two ratios share one constituent, and differ in the second constituent, the two ratios are necessarily different.)

Salome's German and English ratios are more dissimilar than both the adult and Reuben's values. However, her German and English ratios exhibit an inverse pattern, when compared to the monolingual adult values, whereas Reuben's ratios reflect the monolingual adult distribution.

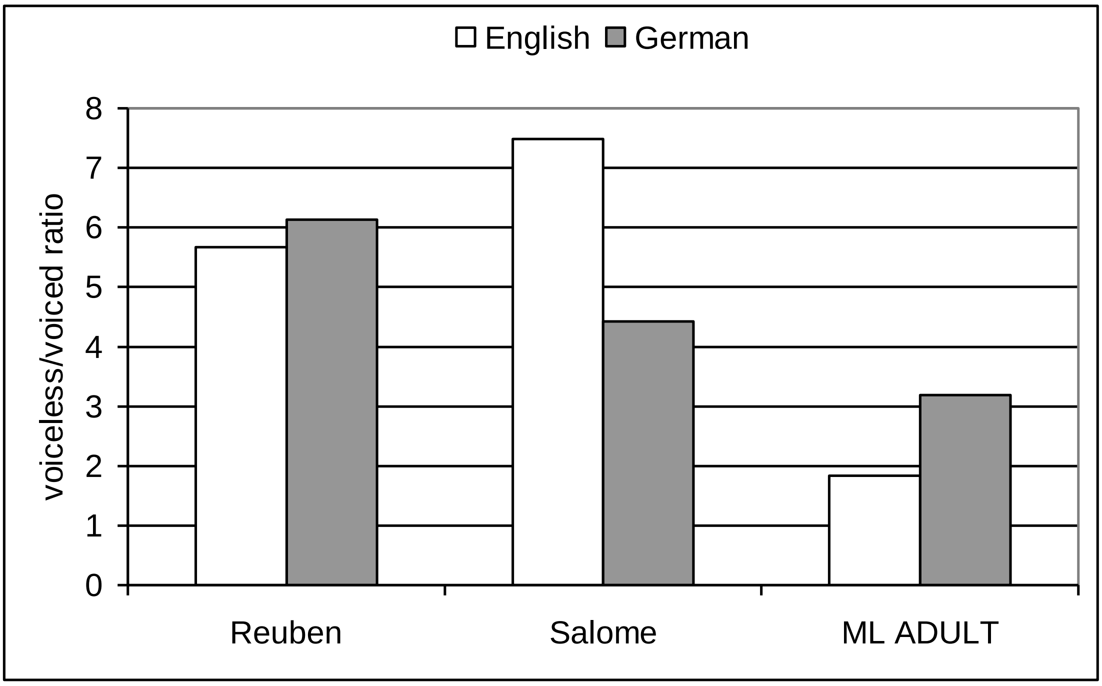


Figure 1: Initial VOT ratio voiceless/voiced

5.2. Intrinsic Vowel Length

5.2.1. Definition

Intrinsic Vowel Length (IVL) refers to the durational differences of vowels in the same phonetic environment, it is thus "determined by the nature of the segment itself" (Lehiste 1970 18). For example, vowel length seems to be connected to tongue height, since close vowels tend to be shorter than open ones (Delattre 1965, Lehiste 1970). This phenomenon can be explained by the longer movement which is required in the case of open vowels to get from and to the articulatory position of the surrounding consonants (Laver 1994).

5.2.2. German and English IVL patterns

In both English and German tense vowels are generally longer than their lax counterparts, e.g. English /ɪ/ *sit* is shorter English /iː/ *seat*. German lax vowels /ɪ ɛ a ɔ ʊ/ are only approximately half as long as the tense vowels, e.g. /iː eː aː oː uː/ (Antoniadis & Strube 1984). In contrast, English tense vowels /iː eː ɑː ɔː uː/ are only approximately one third longer than English lax vowels /ɪ ɛ a ɒ ʊ/ (House 1961).

Table 6: German and English IVL patterns

	English	**German**
Lax vowels (ms)	185[A]	71[B]
Tense vowels (ms)	265 [A]	154[B]
Ratio lax/tense	0.70[A]	0.46[B]

[A]House (1961); [B]Antoniadis & Strube (1984)

5.2.3. Analysis

Vowel duration was measured from the onset of F2 to the offset of F1 (Hewlett, Matthews & Scobbie 1999, Grabe & Low to appear). 122 of Reuben's tokens and 131 of Salome's tokens have been analysed with regard to intrinsic vowel length. The results are summed up in Table 8 and Table 9. The asterisk (*) marks significant differences between the temporal realisations of German tense and English tense vowels as well as German lax and English lax vowels.

5.2.4. Results

Tense Vowels. Reuben and Salome both realise significant durational differences in the production of German and English tense vowels ($p \leq 0.03$). For both children English tense vowels are longer than their German counterparts.

Lax Vowels. Reuben distinguishes between German and English lax vowels ($p \leq 0.06$). His lax English vowels are longer than his German lax vowels. Salome displays the same pattern, however the difference between her German and English lax vowels is not significant.

Table 7: Reuben (9;11) intrinsic vowel duration

	English		**German**	
	lax	**tense**	**lax**	**tense**
No. of tokens	38	27	29	28
Range (ms)	77-235	164-316	87-174	148-282
Standard deviation (ms)	45	47	21	40
Mean duration (ms)	*145.58	*237.19	*132.66	*216.04
Ratio lax/tense	0.61		0.61	

Table 8: Salome (12;5) intrinsic vowel duration

	English		**German**	
	lax	**tense**	**lax**	**tense**
No. of tokens	42	29	25	35
Range (ms)	73-238	132-324	78-186	147-278
Standard deviation (ms)	42	54	29	36
Mean duration (ms)	135.45	*227.55	126.52	*198.94
Ratio lax/tense	0.60		0.64	

In Figure 2 the children's lax/tense ratios are compared to the lax/tense ratios of German and English monolingual adults (see Table 6).

The difference between English and German lax/tense ratios are neutralised in Reuben's productions, although, when looking at the absolute durational values, he significantly differentiates between English and German lax and tense vowels. Reuben's lax/tense ratios (0.61) are somewhere between those of German (0.63) and English (0.55) monolingual speakers.

Salome's lax/tense ratios for English and German vowels are likely to be significantly different, since she distinguishes between English and German tense vowels, but not between English and German lax vowels. Her ratios are greater than the respective monolingual targets. However, the difference between the two languages in adult monolingual speech is imitated in her productions, in so far as her German ratio is also greater than her English ratio.

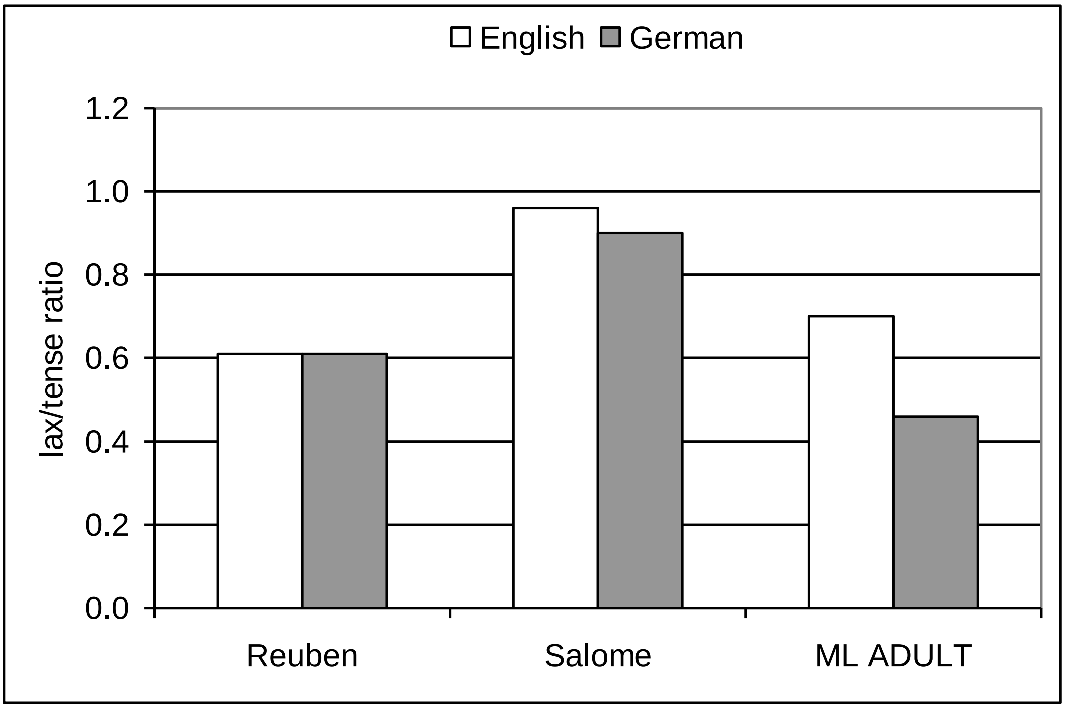


Figure 2: IVL ratio lax/tense

5.3. Extrinsic Vowel Length

5.3.1. Definition

Extrinsic Vowel Length (EVL) refers to the durational differences of vowels in different consonant environments. For example, vowel duration varies depending on the manner of articulation of the following consonant, and on whether that consonant is voiced or voiceless (Lehiste 1970, Chen 1970). Thus, American English vowels are longest before voiced fricatives, and systematically decrease in length when followed by voiced stops, nasal stops, voiceless fricatives, and voiceless stops (Laver 1994).

5.3.2. German and English EVL patterns

In English and German vowels followed by the voiceless stops /p t k/ are shorter than vowels followed by the voiced stops /b d g/ (House 1961, Kohler 1995), e.g. /iː/ is longer in English *seed* than in *seat*. Again the two languages implement slightly different temporal patterns; in English the ratio0 of voiceless/voiced is 0.54 (House 1961), and in German 0.83 (Braunschweiler 1998).

Table 9: German and English EVL patterns

	English	**German**
Preceding voiceless consonants (ms)	170[A]	121[B]
Preceding voiced consonants (ms)	310[A]	146[B]
Ratio voiceless/voiced	0.54 [A]	0.83[B]

[A]House (1961) ; [B]Braunschweiler (1997);

5.3.3. Analysis

Vowel duration was again measured from the onset of F2 to the offset of F1 (Hewlett, Matthews & Scobbie 1999, Grabe & Low to appear). Extrinsic vowel length has been analysed for 146 of Reuben's utterances and 163 of Salome's utterances. The results are presented in Table 12 and Table 14. Significant differences (t-test) between German and English vowel length before voiced consonants and between German and English vowel length preceding voiceless consonants have been marked with an asterisk (*).

5.3.4. Results

Preceding a voiceless consonant. There is a significant difference between Reuben's English and German productions of vowel length before voiceless stops (p[0.02). A *t*-test carried out on Salome's data for vowel length before voiceless stops yielded the result p[0.1. However, it appears that Salome has not acquired extrinsic vowel lengthening even within-languages. The significance of the difference between her productions of vowel length before voiceless and voiceless consonants in English is p[0.3 and in German p[0.06.

Preceding a voiced consonant. No significant differences between the children's productions of vowel length before voiced consonants have been found.

Table 10: Reuben (9;11) extrinsic vowel duration

	English		**German**	
	voiceless	**voiced**	**voiceless**	**voiced**
No. of tokens	41	42	42	21
Range (ms)	77-291	98-355	87-288	121-282
Standard deviation (ms)	59	80	47	54
Mean duration (ms)	191.73*	218.36	166.62*	204.71
Ratio voiceless/voiced	0.88		0.81	

Table 11: Salome (12;5) extrinsic vowel duration

	English		**German**	
	voiceless	**voiced**	**voiceless**	**voiced**
No. of tokens	50	42	39	32
Range (ms)	80-329	90-335	86-304	94-278
Standard deviation (ms)	69	69	54	44
Mean duration (ms)	188	195.6	171.82	190.22
Ratio voiceless/voiced	0.96		0.90	

A comparison of the ratios of vowel length before voiceless stops over vowel length before voiced stops (voiceless/voiced) shows that both children's ratios are greater than those of the monolingual adults (see Table 9). Furthermore, the children both have a greater value for their English ratios than for their German ratios, whereas monolingual adult values are distributed inversely, i.e. the German ratio is greater than the English ratio. It is also apparent that the difference between the languages is not as great in the speech of the subjects than between monolingual English and German speech.

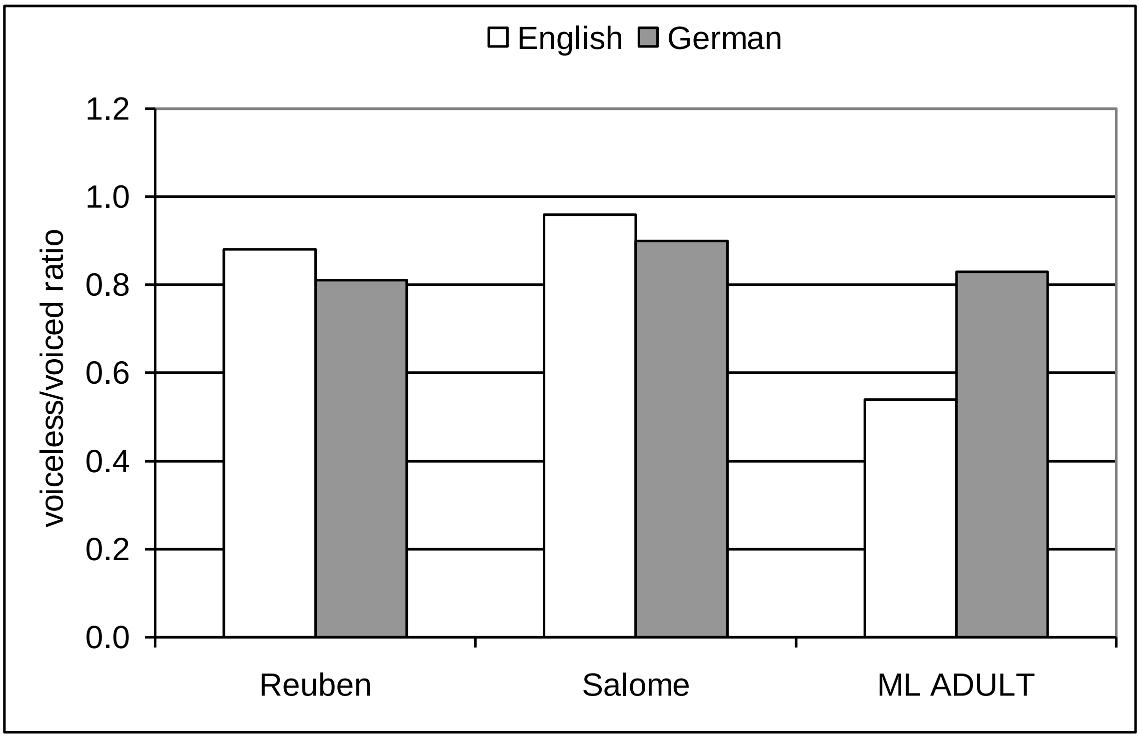


Figure 3: EVL ratio voiceless/voiced

6. Interpretation

The comparison of the children's production with that of monolingual adult productions in both languages has shown that neither child has consistently reproduced the temporal patterns of monolingual adult speech. This indicates that bilinguals cannot be viewed as a simple compound of two monolingual speakers.

However, it is necessary at this point to draw attention to the fact that a comparison of the children's data with the presented monolingual adult patterns has to be undertaken with some caution. The reasons for this are as follows. First, the values for English and German in the different categories have not been taken from sources, which investigated speech timing of both languages within the same study, and thus have not always used the same measurement criteria or linguistic material. Secondly, different studies of the temporal patterns in the same language have produced different values, e.g. the voiceless/voiced ratio of intrinsic vowel length based on the data in House (1961) is 0.70, whereas Hakkarainen (1995) gives a value of 0.55. Apart from differences in analysis procedures, these discrepancies are due to different accents being investigated in each of the studies. Thirdly, the monolingual data which have been used above, possibly reflect a different temporal pattern, than that to which the children were exposed to, since temporal patterns vary between different accents of the same language. To make the monolingual adult data more comparable to the children's, it would have to be collected and analysed within the same study. Moreover, if this data is gathered from the subject's parents this makes it possible to get an insight into the linguistic input the children were exposed to. (This paper only reports the results of a pilot study to a larger project, which will also examine the parents' speech.)

Table 12 gives an overview of the significant differences between German and English in the children's productions of temporal patterns. The symbol "+" in a box means that *t*-test results have shown there to be a significant difference between the child's German and English productions in that category, correspondingly "–"denotes differences which are not significant. There are significant differences between Reuben's productions of English and German VOT for voiceless stops, vowel length before voiceless stops, and the duration of lax and tense vowels. Salome realises significant differences between her productions of English and German VOT of voiceless and voiced stops and the length of her English and German tense vowels.

Table 12: Summary of the significant differences between German and English in the children's productions of temporal patterns (+ significant, – not significant)

	Reuben (9;11)		**Salome (12;5)**	
	voiceless lax	**voiced tense**	**voiceless lax**	**voiced tense**
VOT (voiceless vs. voiced)	+	–	+	+
IVL (lax vs. tense)	+	+	–	+
EVL (voiceless vs. voiced)	+	–	–	–

It can be seen that the children are distinguishing between the temporal patterns of English and German, although (a) neither child reproduces the respective monolingual target, and (b) there is a great deal of variation between the children. The fact that the children to some degree reproduce fine-grained language-specific temporal patterns, reinforces the view that the acquisition of specific phonetic features is an integral part of language acquisition.

The results support the view that the two languages are represented by separate systems, since different timing patterns are applied for each language. The fact that these patterns are dissimilar to the monolingual adult patterns points to some sort of reciprocal influence between the two systems. It should always be kept in mind, however, that the term “system” has not yet been clearly defined (Romaine 1995) and is correspondingly vague.

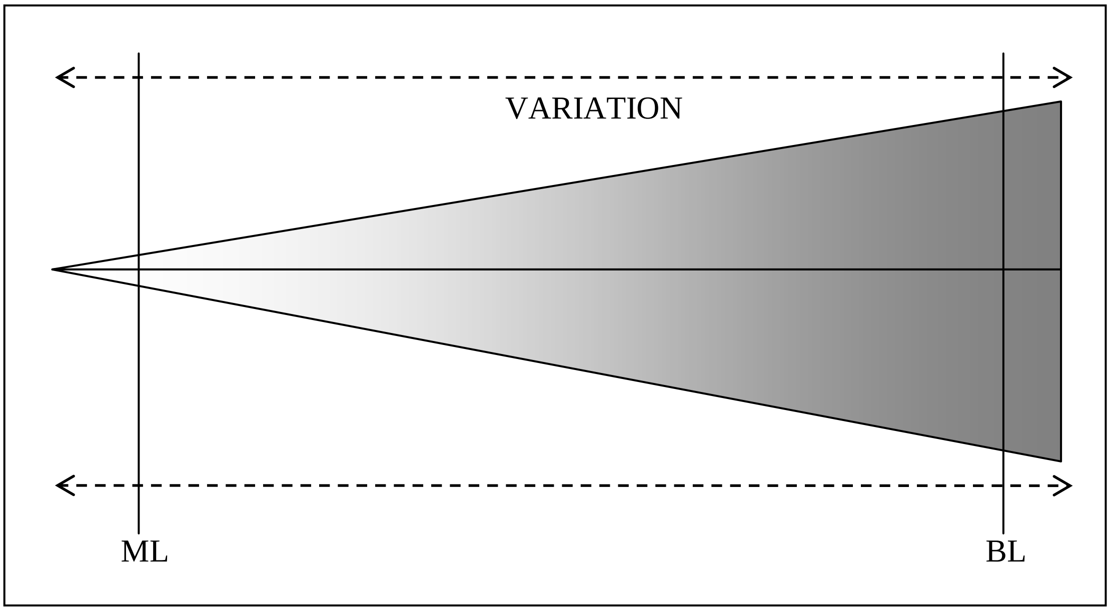
2.3. A theory of bilingual language acquisition

To be able to understand the process of language acquisition, it is necessary to view speech and speakers from the point of view of the child. The infant attempts to construct a valid system from the linguistic input she receives without knowing what the final goal of these attempts is (i.e. language). Locke (1994 279) describes the development of linguistic behaviour as a developmental growth path on which infants are deposited and held to “by a unified force set up jointly by the human genome and perceptual experiences that characterize our species”. Attentional biases (towards the human voice and face), motoric biases (towards certain patterns of movement), and neural specialisations (specialisation in social cognition, grammatical analysis module) help the child to make sense of the input it receives from its environment.

Future monolinguals and bilinguals start the process of language development with the same biological equipment and with the same amount of linguistic knowledge. Two facts have to be remembered. Firstly, even in a monolingual environment linguistic input is not homogenous, e.g. because of different regional accents, sociolects, gender etc. A child brought up in a monolingual environment and a child brought up in a bilingual environment are both surrounded by vocal behaviour which is marked by variation. Secondly, the infant is, at least initially, in no position to judge whether two utterances belong to different languages, dialects, or accents. The prospective bilingual will, thus, treat any linguistic input in the same way as the prospective monolingual child will.

If prospective bilingual infants treat any linguistic input in the same way as the prospective monolingual children, the one-or-two-systems question might be misleading. Let's suppose that monolingual and bilingual language acquisition are a quasi continuum with increasing variation of the linguistic input (see Figure 4), with monolingual language (ML) which contains comparatively small amounts of variation on the one end and further on bilingual language (BL) which is marked by a larger amount of variation. The continuum, for example, starts at one end with language that varies with regard to register, followed by language which varies with regard to register and regional accent, moving on to bidialectal speech, then to bilingualism of two very closely related languages such as German and Yiddish, then to ever less related languages, e.g. English and Spanish, English and Korean, and so on.

Figure 4: Relationship between bilingual (BL) and monolingual (ML) language acquisition: a continuum of increasing structural variation from ML to BL.



It can now be hypothesised that the neural representation of linguistic structures itself might not vary much between monolingual and bilingual subjects, and the difference is merely one of degree, that is amount of neural connections within a network. Current research into bilingual language representation supports a model which permits neural independence for the languages on the one hand and requires some sort of connection between them on the other (Romaine 1995). In the field of monolingual language acquisition Bates (2000) suggests that the neural organisation for language is caused by language learning rather than that language learning is caused by pre-existing neural structures (Chomsky 1986). This emergentist view makes it possible to view bilingualism not as a special form of (monolingual) language acquisition but as an extension of monolingual language acquisition. The question is not any more how an existing language module copes with two or more languages rather than one (for which, as is implicitly assumed, it has been designed), but rather how systematic language variation is represented in the brain. Further research into how sociolinguistic variation is stored in the brain of monolinguals is needed alongside current research into bilingual language representation.

7. Conclusion

There are two conclusions to be drawn from this investigation: First, bilinguals cannot be treated as two monolinguals in one person. Secondly, the results indicate that bilingual children are aware of fine-grained variability in the linguistic input they receive, and are able to produce corresponding patterns which are, however, not identical to the adult monolingual target. This provides further evidence in favour of the dual systems hypothesis. However, the 1/2 systems hypothesis as applied in the field of bilingual language acquisition might be misleading and monolingual and bilingual first language acquisition might result in similar cerebral representation, since within-language variation is likely to be represented by the same neurological structures as between-language variation.

References

Antoniadis, Z. & Strube, H. W. (1984). Untersuchungen zur spezifischen Dauer deutscher Vokale. *Phonetica* **41**. 72-87.

Braunschweiler, N. (1997). Integrated Cues of Voicing and Vowel Length in German: A Production Study. *Language and Speech* **40**. 353-376.

Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B. & Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in French-English bilinguals. *Journal of the Acoustical Society of America* **54**. 421-428.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* **22**. 129-159.

Chomsky, N. (1986). *Knowledge of Language: its Nature, Origin, and Use*. London: Praeger.

Deuchar, M. & Clark, A. (1996). Early bilingual acquisition of the voicing contrast in English and Spanish. *Journal of Phonetics* **24**. 351-365.

Docherty, G. (1992). *The Timing of Voicing in British English Obstruents*. Berlin & New York: Foris Publications.

Fox, A. (2000). *Prosodic Features and Prosodic Structure: The Phonology of Suprasegmentals*. Oxford University Press, Oxford & New York.

Genesee, F. (1989). Early bilingual development, one language or two? *Journal of Child Language* **16**. 161-179.

Grosjean, F. (1998). Studying bilinguals: methodological and conceptual issues. *Bilingualism: Language and Cognition* **1**. 131-149.

Haag, W. K. (1979). An articulatory experiment on voice onset time in German stop consonants. *Phonetica* **36**. 169-181.

Hakkarainen, H. J. (1995). Phonetik des Deutschen. München: Fink.

Hawkins, S. (1979). Temporal co-ordination of consonants in the speech of children: further data. *Journal of Phonetics* **7**. 235-267.

Hazan, V. & Boulakia, G. (1993). Perception and Production of a voicing contrast by French-English bilinguals. *Language and Speech* **36**. 17-38.

House, A. S. (1961). On vowel duration in English. *Journal of the Acoustical Society of America* **33**. 1174-1178.

Khattab, G. (1998). *A Socio-phonetic Study of Arabic-English Bilingual Children*. Unpublished MA Thesis. University of Leeds, UK.

Kohler, K. (1986). Invariance and variability in speech timing: from utterance to segment in German. In: Perkell, J. S. & Klatt, D. H. (eds.). *Invariance and Variability of Speech Processes*. Hillsdale, N.J.: Lawrence Erlbaum. 268-289.

Kohler, K. (1995). *Einführung in die Phonetik des Deutschen*. Berlin, Erich Schmidt.

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

Lisker, L. & Abramson, A. S. (1964). A cross language study of voicing in initial stops: acoustical measurements. *Word* **20**. 384-422.

Locke, J. L. (1995). Development of the capacity for spoken language. In: Flecther, P & MacWhinney, B (eds.). *The Handbook of Child Language.* Oxford: Blackwell. 278-302.

Romaine, S. (1995). *Bilingualism*. Oxford, Blackwell.

Stoel-Gammon, C. & Buder, E. H. (1999). Vowel length, post-vocalic voicing and VOT in the speech of two year olds. *Proceedings of the 14th International Conference of Phonetic Sciences*. 2458-2488.

Vihman, M. M. (1985). Language differentiation by the bilingual infant. *Journal of Child Language* **12**. 297-324.

Volterra, V. & Taeschner, T. (1978). The acquisition and development of language by bilingual children. *Journal of Child Language* **5**. 311-326.

Nicole Whitworth
Department of Linguistics & Phonetics
University of Leeds
LS2 9JT

lnpnwh@leeds.ac.uk