SOCIOLOLINGUISTIC COMPETENCE AND THE BILINGUAL'S ADOPTION OF PHONETIC VARIANTS: AUDITORY AND INSTRUMENTAL DATA FROM ENGLISH-ARABIC BILINGUALS

by

Ghada Khattab

Submitted in accordance with the requirements for the degree of
Doctor of Philosophy

University of Leeds

Department of Linguistics and Phonetics

September 2002

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others

PUBLICATIONS FROM THE THESIS

Khattab, G. (2002). VOT production in English and Arabic bilingual and monolingual children. In D.B. Parkinson & E. Benmamoun (eds.). *Perspectives on Arabic Linguistics XIII-XIV* (pp. 1-38). Amsterdam: John Benjamins.

Khattab, G. (2002). /l/ production in English-Arabic bilingual speakers. *International Journal of Bilingualism*, **6** (**3**): 335-353.

DEDICATION

I dedicate this work to my mum and to Paul.

ACKNOWLEDGMENTS

First, I would like to thank my parents for encouraging my decision to pursue my studies abroad. Their emotional and financial support throughout the last five years was unconditional, and without it I would never have met the demands of the course. Second, I would like to thank the members of staff in the Department of Linguistics and Phonetics at the University of Leeds for seeing the potential in me while I was studying for my MA degree and arousing my interest in pursuing my studies at the doctoral level. Diane Nelson, in particular, has not only played the role of my lecturer, but also that of my advisor and my friend. Her guidance and encouragement have had a huge positive impact on both my academic and personal life. The same is true about Dominic Watt, who was always available for expert help and advice. A lot of the coffee and lunch breaks that we shared consisted of stimulating discussions that later developed into concrete ideas and found their way into the dissertation.

Special thanks are due to my supervisor, Barry Heselwood, whose knowledge of Arabic linguistics, among other areas, is impressive and inspiring. His friendliness, encouragement, and invaluable comments are much appreciated. He was always available for advice, especially at times when I encountered problems during the course of the investigation. I would also like to thank the members of the Phonetics and Phonology research group at the University of York, whose meetings I regularly attended and often used as a platform to discuss my current work. I am also grateful to many experts in the field who I met at conferences and other Academic events, and whose input has positively changed the course of the investigation in many ways. These include Margaret Deuchar, Gerry Docherty, Jane Stuart Smith, Jim Scobbie, and Marilyn Vihman.

My deepest appreciation and gratitude is to Paul Foulkes, who supported me during every single step of the preparation of this work. He has been my inspiration and my shoulder to cry on, and his encouragement, understanding, and enduring patience are unequaled. I dedicate this work to him and to my mum.

ABSTRACT

This study is an auditory and acoustic investigation of the speech production patterns developed by English-Arabic bilingual children. The subjects are three Lebanese children aged five, seven and ten, all born and raised in Yorkshire, England. Monolingual friends of the same age were chosen as controls, and the parents of all bilingual and monolingual children were also taped to obtain a detailed assessment of the sound patterns available in the subjects' environment. The study addresses the question of interaction between the bilingual's phonological systems by calling for a refinement of the notion of a 'phonological system' using insights from recent phonetic and sociolinguistic work on variability in speech (e.g. Docherty, Foulkes, Tillotson, & Watt, 2002; Docherty & Foulkes, 2000; Local, 1983; Pisoni, 1997; Roberts, 1997; Scobbie, 2002). The variables under study include /l/, /r/, and VOT production. These were chosen due to the existence of different patterns in their production in English and Arabic that vary according to contextual and dialectal factors. Data were collected using a variety of picture-naming, story-telling, and free-play activities for the children, and reading lists, story-telling, and interviews for the adults. To control for language mode (Grosjean, 1998), the bilinguals were recorded in different language sessions with different interviewers.

Results for the monolingual children and adults in this study underline the importance of including controls in any study of bilingual speech development for a better interpretation of the bilinguals' patterns. Input from the adults proved highly variable and at times conflicted with published patterns normally found in the literature for the variables under study. Results for the bilinguals show that they have developed separate sociolinguistically-appropriate production patterns for each of their languages that are on the whole similar to those of monolinguals but that also reflect the bilinguals' rich socio-phonetic repertoire. The interaction between the bilinguals' languages is mainly restricted to the bilingual mode and is a sign of their developing sociolinguistic competence.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Following the discovery of a seven million year old skull:

'... Toumai enlightens us because it disproves our theories, and shows us again that we too often assume that the data available to us contain all we need to answer our questions. So one lesson of this great discovery — an old lesson, but always worth relearning — is that we should never assume a lack of evidence means the evidence doesn't exist, or that the evidence we have is at all definitive. In science, a bird in the hand is not worth two in the bush. As we try to figure out the meaning of Toumai, we should try to remember never to assume we've learned enough to answer our questions.'

DANIEL E. LIEBERMAN, *New York Times*, July 14, 2002.

Upending the Expectations of Science

CHAPTER ONE

Introduction and Background

1.0 Introduction

This study is an investigation of the speech production patterns developed by three English-Arabic bilingual children aged five, seven, and ten. All three subjects were born and bred in Yorkshire, England, while their parents are native speakers of Lebanese Arabic who had been living in the UK for 10-15 years at the time of the investigation. The study draws on insights from three different but interconnected disciplines: childhood bilingualism, sociolinguistics, and phonetics and phonology.

In this chapter, I discuss the general issues in these areas. Section 1.1 presents some of the definitions of childhood bilingualism that are available in the literature, while Sections 1.2 and 1.3 address the main question that has occupied bilingual research for the last three decades, that is whether bilingual children start with one system or two for their languages in the early stages of development. Then, using theoretical considerations in bilingual and monolingual phonological acquisition (Sections 1.3-1.4), the one-or-two-system question is reassessed and declared problematic due to the simplistic manner in which the notion of 'system' is normally dealt with in the literature. An attempt is then made to refine this notion from the view point of what it means to acquire a 'phonological system'. This is achieved through assessment of recent research on early perception and production abilities in the child (Sections 1.5.1-1.5.2), developmental patterns and individual differences in speech perception/production (Section 1.5.3), sociolinguistic variability in the input (Sections 1.5.4-1.5.5) and the role of sociolinguistic variability in the acquisition of sociolinguistic competence in the child (Section 1.5.6). The implications of recent phonetic and sociolinguistic research for bilingual phonological acquisition are then discussed in Section 1.5.7, along with other important methodological issues related to bilingual research (Section 1.6). The aim is to set the stage for the methodology of the current study, which will be presented in Chapter Two.

The study is innovative in many ways. First, it deals with two languages that have rarely been studied in combination in bilingual phonological acquisition. Second, it adopts a different stance on what is meant by a phonological system, by virtue of a grounding in aspects of sociolinguistics. Recent phonetic and sociolinguistic work on monolingual acquisition has argued that there is no simple stable phonological model that any child is exposed to (e.g. Docherty et al, 2002; Foulkes, Docherty and Watt, 1999; Local, 1983; Roberts, 1997). Instead, there may be considerable variability in the input that a child receives that is not only phonological/allophonic, but may also be linked to social characterisation of the speaker (sex, age, social class, etc.) and to speaking style. Such factors create variability in the input for any child in any community, and must therefore be mastered alongside aspects of the contrastive phonology.

The bilingual child faces an added degree of variability by being regularly exposed to input that may vary between standard, non-standard, and non-native varieties for either language, especially if the parents are non-native speakers of the community language. The varied input that the bilinguals receive in each language and from different speakers is bound to influence their phonological development. Patterns that are specific to one language might be used in the production of the other language, and some patterns may or may not be acquired depending on individual differences and the amount of input that is crucial for their acquisition. While language interaction is characteristic of bilingual speech, the amount of control that bilinguals have in trying to keep their languages separate depends on controversial factors such as age, dominance, and proficiency. What is important, though, is to interpret the bilinguals' linguistic behaviour in context, i.e. in relation to whether or not they show control over the production of these patterns depending on the identity and expectation of their interlocutor.

In order to take variability into account, the study adopts a different methodology in that it does not only rely on published accounts of production patterns in either language. Very few studies have examined the actual phonological input that the bilinguals receive in order to verify its compatibility with published sources. Published sources may report patterns that are found in standard descriptions or in varieties other than the ones examined by the researchers. These reports may be outdated, or may have been obtained using different methodologies which have bearings on the outcome of the patterns observed. Since it is likely that the bilinguals' social network has an influence on their linguistic choices, monolingual English friends of the bilingual children were also recorded for the project, along with monolingual Arabic controls and the parents of all bilinguals and monolinguals.

The consonants that were chosen for investigation vary in their production patterns not only between the two languages in question, but also within each language depending on dialectal and sociolinguistic factors. For this reason, the study combines detailed phonetic analysis using instrumental techniques and sociolinguistic analysis based on principles of regional and social variation in order to get a more detailed view of the phonological targets that the bilingual children in this study are aiming for. The main aim is to find out whether the bilinguals have acquired separate sociolinguistically appropriate production patterns for these targets in each language. In the case of influence between the two languages, the aim is to find out whether language interaction is under the bilingual's control and whether it varies depending on the bilingual's language mode (Grosjean, 1998).

1.1 The many definitions of bilingualism

Childhood bilingualism is the area of language acquisition concerned with the simultaneous or sequential acquisition of two languages by children. Simultaneous acquisition refers to children who receive input from both languages from birth or before their third birthday, while successive or sequential acquisition takes place where input from a second language is received after the third birthday (Lyon, 1996: 47). This, however, constitutes only one way of classifying young bilinguals; different researchers have used different terms and different age limits when describing types of bilingualism. De Houwer (1995: 223), for instance, suggests the term 'Bilingual First Language' (BFL) acquirers for situations where the child is regularly exposed to two languages within the first month of birth. She argues that situations where regular exposure to a second language occurs later than one month after birth and before the age of two should be categorised as 'Bilingual Second Language Acquisition'.

Many neurolinguistic and psychological studies have also distinguished between various types of bilingualism. One of the earliest distinctions of types of bilingualism was made by Weinreich (1953), who distinguished between three categories of bilingualism: compact, coordinated, or subordinated bilingualism. The expression 'compact bilingual' refers to an individual who has learnt the two languages simultaneously before the sixth year, normally because they were each spoken by one of the parents. A 'coordinated bilingual' has learnt the second language before puberty, within or outside the family, for example because the child moved to another foreign country with the family. A 'subordinate bilingual' has one language as the mother tongue and uses the second language as moderator of the first language. In this type of bilingualism, subjects think of what they want to express in their first language first and then translate it into their second language.

Other terms includes 'primary bilinguals', which refers to the acquisition of both languages in natural contexts and usually before the age of three, and 'secondary bilinguals', which refers to cases where one of the languages is acquired after the age of three (Hoffmann, 1991: 19; Lyon, 1996: 48). Similar comparisons are drawn using the terms 'early bilingualism', which refers to early acquisition (in infancy) of the two languages, and 'late bilingualism', where the second language is acquired much later than the mother tongue (though there is no agreed upon age limit between early and late bilingualism). Other definitions attempt to describe the degree of competence in the two languages; for instance, a 'balanced bilingual' is a subject who has mastered two languages to the same extent, whereas a 'dominant bilingual' is a subject who is more fluent in one language than the other (Fabbro, 1999: 107).

In any specific case, the individual circumstances surrounding the language acquisition of every child are different and do not necessarily fit in any of the categories described in the literature. As Grosjean (1982; 1995) notes, 'the bilingual is not two monolinguals in one person'; rather bilingual individuals have differentiated needs for their two languages or attribute them to different social/emotional functions (what a language is used for, with whom, where, etc.). Thus, they do not necessarily have to develop perfect knowledge, nor the same level of competence and/or performance in both languages.

1.2 One or two systems?

The growing number of bilingual speakers all over the world (Fabbro, 1999: 103; Grosjean, 1982; Holmes, 1992: 79; Tucker, 1998) has recently been accompanied by a parallel growth of interest in the study of bilingual children's language development and in crosslinguistic studies of language acquisition. The main question that has occupied researchers since the 1970s and 1980s is whether bilingual children (i) start by mixing both systems of the two languages and later separate them during the chain of their development (Gradual differentiation theory) (e.g. Leopold, 1970; McLaughlin, 1984; Redlinger & Park, 1980; Schnitzer & Krasinski, 1994; Swain, 1972; Volterra & Taschner, 1978); or (ii) separate the linguistic systems of their two languages from the beginning of their language development (Separate Development Theory) (e.g. De Houwer; 1990; Deuchar & Quay, 1999; Genesee, 1989; Genesee, Nicoladis, & Paradis, 1995; Lindholm & Padilla, 1978; Lanza, 1997; Petitto, Katerelos, Levy, Gauna, Tetreault, & Ferraro, 2001; Schnitzer & Krasinski, 1996).

The question is a very complicated one, as it touches upon unresolved issues in both monolingual and bilingual acquisition. These relate to (i) infant perceptual abilities and their relation to later production; (ii) the cerebral organisation of language(s) in the brain; (iii) the nature of the knowledge (or mental representation) that underlies language performance; and (iv) the influence of the sociolinguistic environment on the development of language(s) in the child. Each of these issues will be dealt with in this chapter, but first, we turn to existing studies that have contributed to the one-or-two-systems debate.

1.2.1 The unitary language system explanation

From a neuro-cognitive perspective, the unitary-system explanation suggests that, during the initial stages of language development, the language faculty is biologically and, therefore fundamentally 'monolingual' (Genesee, 2001: 155). That is, the human brain is neurologically set to acquire only one language at birth and must undergo fundamental neural reorganisation in the face of two (Petitto et al, 2001: 490). Claims for the unitary system with undifferentiated syntactic, lexical, phonological and subsystems have been based on reported findings of children mixing elements (phonological, lexical, morphosyntactic) from their two languages in the same utterance or stretch of conversation. These claims are also supported by a noticeable decrease in mixing with age, although there are exceptions to that (e.g. Vihman, 1985).

Most of the evidence for the unitary system explanation comes from longitudinal case studies of bilingual language development. One of the earliest studies of bilingual development is by Ronjat (1913), who claimed that children originally draw on both languages in an apparently indiscriminate way, and that they neither distinguish between their interlocutors nor make an obvious effort at consistency of language choice within a given utterance. The child is then assumed to eventually sort out the lexicon, phonology, and grammar given continued exposure to both languages, and to surprise observers with his/her pragmatic facility in addressing the right language to the right interlocutor. Similarly, Swain (1977) postulated a 'common storage model' of bilingual development according to which all rules of both languages are initially stored in a common location. A process of differentiation would then tag each rule as being specific to a particular language. Volterra & Taeschner (1978), who studied lexical and syntactic development in their Italian-German bilingual daughters, suggested three stages during which the child gradually becomes bilingual from early infancy: (i) the child has one lexical system which includes words from both languages; (ii) the child distinguishes two different lexicons but applies the same syntactic rules to both languages; (iii) the child has two linguistic codes, differentiated both in lexicon and syntax, but each language is exclusively associated with the person using that language. At the end of this stage (around the age of three), when the tendency to categorise people in terms of their languages decreases, the child is considered to be truly bilingual.

Vihman (1985) also studied lexical and syntactic development in her Estonian-English bilingual son Raivo, and initially suggested a gradual transition from a single lexicon with a few corresponding terms to a dual lexicon in which the smaller proportion of English terms was largely duplicated by Estonian terms. Vihman noted that, while in the early stages Raivo did not appear to be concerned with the difference between language sources, contexts and interlocutors, he showed differentiation from the age of 2;0. However, this was the age around which Raivo's linguistic ability increased in both languages and his attention to the pragmatics of his bilingual situation appeared to have grown. Therefore Vihman observed that, since Raivo initially had lower exposure to

English than Estonian, this might have given the initial impression that he was using one system. Moreover, Raivo knew that both his parents were bilingual, and his mixing habits might have been related to his awareness that his parents spoke both languages. Vihman also stressed the importance of taking comprehension into consideration since Raivo's comprehension appeared to progress rapidly in both English and Estonian contexts well before the development of a wide-ranging productive vocabulary. She therefore acknowledged that two receptive stores might have existed at the earliest stage even if in a rudimentary form and concluded that bilingual children are able, from an early age, to differentiate their two linguistic systems.

1.2.2 Criticisms of the unitary system explanation

Most proponents of the unitary-system hypothesis have been criticised for not presenting or analysing their data by context (see discussions by Genesee, 1989: 166; Genesee, Nicoladis, & Paradis, 1995: 613; Quay, 1995: 385). Therefore it is impossible to determine whether the children are using the repertoire of language items they have acquired to that point in a differentiated way. For example, studies like Volterra & Taeschner's (1978) have given examples of the child's mixing when communicating with her German-speaking mother, without comparing the child's behaviour with, for instance, the Italian-speaking father (the family lived in Italy). In comparison, Goodz (1989) and Lanza (1992) examined bilingual children's use of their two languages with each parent, and both reported that the children they observed code-mixed very little with their parents, and therefore generally used the appropriate language with different interlocutors. Goodz (1989) also found a positive correlation between parental rates of mixing and that of their children.

Similarly, Genesee et al (1995) compared language use by five French-English bilingual children aged between 1;10 and 2;2 with each parent separately and together. Two of them were also taped while playing with a monolingual English-speaking stranger child. All five children, regardless of their language dominance, used more English with their mothers (the mothers are all native-English speakers) and more French with their fathers (the fathers were all native-French speakers), regardless of whether the children were with one or two parents (Genesee et al, 1995: 622). The two children who were also taped playing with monolingual English children used more English-only utterances than French-only or mixed utterances. The fact that these children still produced French while interacting with monolingual English strangers was explained in terms of constraints on their proficiency in English. The role of dominance also showed in the way the children mixed more with the parent who spoke their less dominant language (Genesee et al, 1995: 626).

While Volterra & Taeschner (1978) proposed that bilingual infants in stage one do not have 'translation equivalents', i.e. crosslinguistic synonyms, Quay, (1995) found evidence for these synonyms from a very early stage of lexical development (age 1;1) in her subject, Manuela. Quay attributed the differences in results to the fact that Volterra & Taeschner's (1978) observations were only based on interactions between the children and their German-speaking mother, which might have excluded translation equivalents that the child could have produced in the presence of her Italian-speaking father. Quay's observations, on the other hand, were made from data that included interactions between Manuela and her monolingual Spanish father on the one hand, and her monolingual English grandmother on the other. The availability of the same toys and books in the two language contexts provided strong evidence for bilingual synonyms (Quay, 1995: 385). Similar results were found by Pearson, Fernandez & Oller (1995) for a group of 27 developing English-Spanish bilinguals between the ages of 0;8 and 2;6. The authors also underlined the importance of looking at a large number of children before drawing conclusions about the early lexical development of bilinguals, as there were considerable differences between the children that they observed, with some bilinguals avoiding translation equivalents, and others openly accepting them (Pearson et al, 1995: 364).

While Redlinger & Park (1980) used the decrease in the amount of code-mixing in their child's over time as evidence for gradual differentiation, Genesee (1989: 166) points out that mixing might actually decline because the children are acquiring more complete linguistic repertoires, and do not need to borrow between languages. Similarly, Vihman (1985: 313) points out that mixing might decline with age, and that the child might become aware of adult standards of behaviour and show his/her ability to meet them. This can be interpreted as a sign of the child's developing a sociolinguistic competence.

Another factor that needs to be seriously considered before 'condemning bilingual children to linguistic chaos' (Bialystok, 2001: 114) in the earliest stages of language acquisition is an analysis of what they actually hear. Though the most common model described in the literature for raising bilingual children is the one parent, one language arrangement, Bialystok wonders how realistic this model is. For instance, according to parental reports in Goodz (1989) the young bilinguals who were being investigated were being raised in this way. Parents assured researchers that they were careful about honouring the household rules of linguistic choice and that their speech to the child was 'pure' and 'unadulterated'. However, in recorded observations from the home, Goodz found otherwise. Parents did mix languages, and children's own integrated utterances may well have been a reflection of the language model for them at home.

Similar results were obtained by Hiroko (1998), who investigated mixing patterns among Japanese-English bilingual families living in the US. The parents claimed a strict use of the one-parent-one-language approach. During the study, however, it became clear that the strict use of Japanese by Japanese-speaking parent was not maintained in interactions with their children. The primary goal of the parents seemed to encourage language behaviour irrespective of its form. In monolingual families, this motivation is shown by the parents' tolerance of a variety of linguistic errors. In Hiroko's (1998: 337) study, all parents used more English in response to a non-Japanese utterance by their child than in any other circumstances. A parental code-switch to English almost always led to the child using English (more often than not triggered by the child's use of English to begin with). Hiroko (1998) concludes that the desire to maintain the minority language is difficult if the parents signal to the child that it is acceptable to use English. However, the author notes that children may not be failing to learn the minority language just because they do not speak it very often, as long as they continue to be exposed to the minority language at home. They are actively listening to the language and storing up information on it. When the child becomes motivated enough to use the language, many parents may be surprised at the rapid progress that is made. This was true for one of the subjects who were recorded for the current study, as her efforts to speak Arabic increased tremendously during a visit to relatives in the Lebanon, whereas her parents' efforts to encourage her to speak Arabic in the UK are normally in vain.

1.2.3 Explanations for mixing

Alternative explanations for mixing can be classified into two broad categories: those that are input-based and those that are proficiency-based (Genesee, 2001: 156). According to input-based explanations, bilingual children code-mix because of the input addressed to them by others. In other words, children who are exposed to extensive code-mixing by older siblings, parents, and other adults code-mix more than children exposed to less mixed input, especially if code-mixing is encouraged and accepted by the adults (e.g. Fantini, 1985; Redlinger & Park, 1980). Such an explanation seems plausible, since children would be showing signs of having acquired the patterns and forms of code-mixing that occur in communities as part of their language socialisation. However, Genesee (2001: 156) notes there has not always been statistical evidence in support of this hypothesis (e.g. Genesee et al, 1995; Lanza, 1992; Deuchar & Quay, 2000).

According to the proficiency hypothesis, young bilinguals code-mix to fill gaps in their language proficiency. Evidence comes from findings that bilingual children mix more when using their less proficient language (Genesee, Nicoladis, & Paradis, 1995). Children may lack appropriate lexical items in one language but have them in the other language, and therefore borrow from one language for use in the other (Deuchar & Quay, 2000; Genesee et al, 1995; Fantini, 1978; Lindholm & Padilla, 1978; Volterra &

Taeschner, 1978, Redlinger & Park, 1980). Even fully proficient adult bilinguals do this when they experience a temporary lack of memory or when an appropriate word or expression does not exist in the language they are speaking (Genesee, 2001: 157). Another reason for mixing might be due to a restricted use of specific lexical items by the child from the language that was first or most frequently used to label these items, resulting in the child identifying the referents with the lexical items of that language regardless of the linguistic context. When the mixes occur, the structural consistency of the utterance is maintained (Lindholm & Padilla, 1978).

A third explanation of proficiency-based mixing has been offered in terms of structural linguistic factors. Vihman (1985), for instance, notes that her son used English function words in otherwise Estonian utterances because the English words were 'simpler and more salient' than the corresponding Estonian words. This notion of saliency was also reiterated by Lindholm & Padilla (1978), who noted that bilingual children employ language mixes either when they lack the lexical entry in the appropriate language or when the mixed entry is more salient (in this case they give the e.g. of 'y' in Spanish, which appeared to be more salient than English 'and' and was therefore used more frequently by the children). Moreover, the language system of the child may be incomplete and may not include all the grammatical devices needed to express certain meaning. If a device from the other language is available and serves the same purpose, it might be used temporarily (Genesee, 1989: 168). Developing bilingual children can therefore be seen to be using whatever grammatical devices they have in their repertoire or whatever devices they are able to use given their current language ability.

Altogether, the explanations offered for bilingual code-mixing indicate that mixing should not reflect an inability of the language faculty to develop two different systems during the initial stages of acquisition. As Genesee (1989: 167) points out, some of the explanations for code-mixing can be in fact be interpreted in terms of acquisitional processes that have been identified in monolingual acquisition. For instance, mixing due to lexical borrowing can be viewed as overextensions of the type observed in monolingual acquisition (e.g. using the same vocabulary item for several referents), with bilingual children extending within and across languages and monolingual children extending within languages only.

1.2.4 The Separate Development explanation

Contrary to earlier hypotheses, it is generally agreed now that the languages of the bilingual child are represented in underlyingly differentiated ways at least from the beginning of early language production, and possibly earlier (Genesee, 2001: 158). Numerous studies have found evidence that bilingual children can use their developing languages differentially and appropriately with different interlocutors from the earliest stages of productive language use (e.g. De Houwer; 1990; Genesee, 1989; Genesee et al, 1995; Lindholm & Padilla, 1978; Lanza, 1997; Meisel, 1989; Petitto et al, 2001). As mentioned in Section 1.2.2, evidence comes from observing that children as young as 1;0 use more of their mothers' language with their mothers and their fathers' language with their fathers (e.g. De Houwer, 1983; Genesee et al, 1995; Petitto et al, 2001) or use the appropriate language with monolingual strangers (Petitto et al, 2001).

These findings are significant because they indicate that pragmatic differentiation is evident from a very early productive stage (one-word stage), and that bilingual children have the cognitive capacity to identify important communicative characteristics of their interlocutors and to respond appropriately (Genesee, 2001: 156). De Houwer (1990) points out that children who mix languages may be making sociolinguistic errors of language choice but they are not necessarily making psycholinguistic errors. Moreover, studies on syntactic development show that bilingual children from as early as two years of age use language-specific and different syntactic constructions when addressing interlocutors who speak different languages (e.g. De Houwer, 1990; Ingram, 1981, 1982; Paradis & Genesee, 1996).

While each of the above-mentioned studies (along with many others in the field) provides strong evidence for the separate development hypothesis, Petitto et al's (2001) study is of particular importance, due to the original methodology which consisted of examining bilingual acquisition from the earliest utterances (ages 1;0-3;6) across two modalities, spoken and signed. The authors compared simultaneous bilingual language acquisition in two groups of children: one group was acquiring a language in the spoken modality, French, and a language in the signed modality, Langue des Signes Québecoises (LSQ), and another group was acquiring two languages in the spoken modality, French and English. The unique situation of the French-LSQ children allowed the investigators to identify all utterances as belonging to one or the other language of the bilinguals (as opposed to the frequent difficulty of classifying early utterances in a child acquiring two spoken languages). Moreover, since the spoken and signed modalities are physically different and could therefore be used at the same time, the authors investigated the possibility of simultaneous language-mixing by the children, and its implications for the bilingual's ability to establish stable and independent language representations.

The results for the children acquiring a spoken and a signed language were very similar to the French-English subjects, and showed that the young bilinguals were not delayed in the achievement of the early milestones in each of their respective languages, but rather displayed patterns that were very similar to those of monolingual children acquiring each of the languages in question (Petitto et al, 2001: 469). More interestingly, both groups of children produced translation equivalents from an early age, which showed their awareness of the two languages, and showed high interlocutor sensitivity by making language choices that were related to their addressee. For example, more French would be used with a native-French parent or a monolingual French stranger, and more English or LSQ would be used with a native-English/LSQ parent or a monolingual English/LSQ stranger (Petitto et al, 2001: 479). All children also produced language mixing, but each child's rate of mixing was found to be directly related to their parents' rate of mixing and their language preference (following influence from parent, sibling, carer, nursery, etc.), as well as the need for 'guest words', particular words they are only used to producing in one of their languages. Simultaneous language-mixing in the LSQ-French children mainly contained semantically congruent mixes (the signs and the words had the same meaning), which showed that the mixes were semantically appropriate. As for the semantically incongruent mixes (the signs and words had a different meaning), these preserved the correct syntactic order in the grammar of each respective language and complemented each other to produce a cohesive whole (Petitto et al, 2001: 488).

Altogether, the results from Petitto et al (2001) showed that there was no initial confusion in the children's production even at the earliest stages, and that language mixing was systematic and principled from the time the children started the language acquisition process in production (age 0;10). According to the authors, such results provide evidence for distinct representations of the input languages in bilingual infants from the very first steps in the language acquisition process.

1.3 Bilingual phonological acquisition

Until recently, research on phonological acquisition had received less attention in the field of bilingualism in comparison with investigations of other areas of the grammar (e.g. De Houwer, 1998: 256). One of the reasons behind the scarcity of research in this area is the difficulty of interpreting children's early stages of sound production, even in the case of monolingual acquisition (e.g. de Boysson-Bardies & Vihman, 1991; Eilers, Oller, & Lavoie, 1985). When researching bilingual phonological acquisition, the task is more complex because two sound systems are involved, and a larger number of sound characteristics need to be examined as they are acquired and produced. Watson (1995) notes that there are already so many theoretical choices when it comes to analysing the adult sound system of a single language that it is so difficult to start to find a conceptual framework which will permit bilingual acquisition to be investigated.

As with other areas of bilingual acquisition, the issue of whether a bilingual child starts with one phonological system or two at the onset of language development has produced mixed results, in part due to differences in the methodologies used in studying bilinguals, but also due to a problem that is inherent in the question itself (discussed in Section 1.4). Opinions are divided as to whether the child starts with (i) a single sound system (e.g. Contreras & Saporta; 1970; Leopold, 1970; Schnitzer & Krasinski, 1994); (ii) two different systems (e.g. Ingram, 1981; Paradis, 1996, Schnitzer & Krasinski, 1996); (iii) no system (e.g. Deuchar & Quay, 2000; Johnson & Lancaster, 1998; Major, 1977); or (iv) two independent but non-autonomous systems (e.g. Holm & Dodd, 1999; Paradis, 2001; Watson, 1991).

1.3.1 The one-system option

Leopold's (1970) diary of his daughter's bilingual development in English and German is a common starting point for studies of bilingual development, particularly his claim that Hildegard started out with a unified language system. Analysis of Hildegard's phonological development from diary notes consisted of a segmental inventory, a substitution analysis, a description of the development over seven years, and a discussion of selected processes (assimilation, metathesis, etc.). Leopold (1970: 206) concluded that, in the first two years, Hildegard was still trying to 'weld the two linguistic systems into one unit'.

In another study, Contreras & Saporta (1970) investigated the phonological development of a child acquiring American English and South American Spanish simultaneously from ages 1;0 to 1;7. The authors also argued that the child had an initial single system, based on the suggestion that there is a single set of phonemes that have various allophones in the child's speech. The distribution of allophones in their data is not identical for English versus Spanish words, but all words are treated as belonging to the same set, to which tests for complementary distribution and phonetic similarity are applied. For instance, Contreras & Saporta argue that [ẽ] and [e] are allophones of the same phoneme /e/ on the grounds of phonetic similarity and complementary distribution (since [ẽ] occurs before a nasal in American English). However, [ẽ] occurs only in an English word ('man') and [e] in a Spanish word (*pepe*), so that they could equally well have argued from this data that /e/ was a phoneme of Spanish only, and /ẽ/ of English.

Schnitzer & Krasinski (1994) studied the development of their Spanish-English bilingual son Fernando from age 1;1 through 3;9. Their methodology consisted of narrow transcriptions of a diary, along with half-hour videotapes which were made twice a month. The authors followed a phone-matrix analytical tool, which is based on the 'phone-tree' as described by Ferguson & Farwell (1975), and which consists of considering all the sounds which a given child uses in a given position of a given word at a given stage of language acquisition. The target segments that were examined were described as 'roughly equivalent to structuralist allophones or to generative systematic phonetic units, essentially what a non-native speaker would need to learn in order to attain a native-sounding pronunciation' (Schnitzer & Krasinski; 1994: 591). The authors admit ignoring certain detailed phonetic differences between the two languages, such as the difference between Spanish [a] and English [ɑ], alveolar [t], [d] in English and dental [t̪], [d̪] in Spanish, vowel lengthening in English, and other 'free stylistic or sociolinguistically conditioned variants' (Schnitzer & Krasinski, 1994: 591).

Schnitzer & Krasinski (1994: 614) concluded that Fernando started with a unitary system for both English and Spanish consonants at about age 1;1, for a short period prior to the establishment of separated systems by 2;7. However, what was described as a unitary system did not take into account normal developmental patterns. For instance, [w], [l], and [d], were found as belonging to a single system between age 1;1 and 2;3, until English [ɹ] was introduced. But English [ɹ] is normally acquired late in monolingual acquisition (see Chapter Four). Similarly, a single lateral is reported as being used in all positions in both languages, with zero replacing a previously used dark [ɫ] (by zero, the authors might have included /l/ vocalisation). But dark [ɫ] is also acquired late (see Chapter Three), and therefore the fact that the child did not produce it in English until age 2;3 might be due to developmental reasons. The authors do comment on the fact that Fernando produced [w] and [l] for Spanish [ɾ] until 2;8, when English [ɹ] alternated with [w] and [ɹʷ], and therefore [w] ceased to be used in Spanish. Again these are developmental factors that have been noted elsewhere. Dark [ɫ] was reintroduced in a stable way at age 2;3, indicative of two systems (Schnitzer & Krasinski; 1994: 615). Vowels were not found to go through a unitary stage, but are described as displaying a 'chaotic pattern' due to extensive variation, even at age 2;7 when the consonant systems are separated. The vowels then stabilised at age 2;8 - 2;9.

Schnitzer & Krasinski (1994) proposed five principal stages: (i) the introduction of phonetic type, often not as part of a system; (ii) the establishment of a unitary system; (iii) the establishment of separate systems; (iv) achievement of target values of the adult system; and later interference of one language on the other. In their conclusion, however, the answer to the one-or-two system was still inconclusive, as Schnitzer & Krasinski (1994: 619) noted that one might be imposing the idea of a system (in terms of mental representation) upon the emerging phonological production, when such emergence is at the mercy of articulatory maturation, and is systematic only in the sense that some sets of sounds are more easily produced first than others. This proved true in their second study, which we now turn to.

1.3.2 The two-system option

In another longitudinal study conducted by Schnitzer & Krasinski (1996) on an older sibling of Fernando, Zevio, the authors found evidence for initial independent development of both vocalic and consonantal segments in the child's production in both languages. Of particular interest is the development of liquids in the child's production, whereby [l], which was introduced stably in Spanish at 1;6 and English at 2;2, never appeared for English [ɫ], except briefly at 3;1, after which [ɫ] was established stably (Schnitzer & Krasinski, 1996: 556). Similarly, Spanish [r] had a separate development from English [ɹ], as the former had its own developmental patterns, being substituted by [hɾ], [hɾ̥], [ɾ̥], and [ɾ̥], all possible variants of [r] in Puerto Rican Spanish. English [ɹ] displayed very different patterns from Spanish [ɾ], as it never occurred as [j] or [y] (as did the Spanish [ɾ]), and was instead frequently labialised, sometimes resulting in a [w] realisation (Schnitzer & Krasinski; 1996: 556). Moreover, the authors found no evidence of later interference of one language on the other in Zevio's production, and concluded that Zevio's development was very much like that of two monolingual children.

In trying to explain the differences in the results found for the two brothers, Schnitzer & Krasinski (1996: 557) note that transcription for Zevio did not begin till age 1;6 (as opposed to 1;1 for Fernando) because there was not enough to transcribe. Fernando completed his first fifty words at 1;4, four months earlier than Zevio, and his first 100 words at 1;7, three months earlier than Zevio. Therefore, Zevio's relatively slower language development may have been due to a strategy of learning two separate systems from the beginning, rather than a single system as a first stage. In many cases, Zevio was described as not producing a phone until he was ready to incorporate it into a system, and therefore he used correct and stable segments from the outset, as opposed to his brother who had unstable productions at first. Zevio's results triggered a revision of the conclusions that were drawn for his brother Fernando (Schnitzer & Krasinski, 1994), as in the second study (Schnitzer & Krasinski, 1996) the authors attributed the apparent initial tendency for Fernando to use a single consonant system to the fact that he was articulatorily incapable of making distinctions necessary to distinguish the two languages.

In another study, Paradis (1996) reanalysed Leopold's (1970) data of Hildegard's productions using a syllabic level of analysis and found evidence for two separate systems. Paradis found that Hildegard produced more reduplication in English than in German, and that this could not be attributed to differences in the input from English and German. This led him to conclude that prosodic development reveals more than analysing phonetic inventories at such an early stage of development, and that research on phonological differentiation should not be limited to segmental aspects of speech.

In his study of a two-year-old child acquiring English and Italian, Ingram (1981-1982) took the opposite approach from Contreras & Saporta (1970) with regards to their assumption that a single set of phonemes were being used for both languages. Ingram analysed the phonology of the child's two lexicons separately, and examined phonetic inventories of the two languages, along with the proportions of monosyllables, closed syllables and reduplications. His results showed that the child's phonological output for English and Italian was highly influenced by the phonological form of the adult models. There was therefore evidence for two phonological systems in the sense that there were specific tendencies in the output to help identify words as belonging to one lexicon or the other.

1.3.3 The no-initial-system option

Recently, some researchers have started to wonder whether it is appropriate to raise the question of one versus two systems in relation to developing bilinguals under the age of two years (e.g. De Houwer, 1995: 231-5). For instance, Major (1977) approached the study of the phonological differentiation of an English-Portuguese bilingual by focusing on phonological processes. He found that similar phonological processes affected segments from the two languages up to the age of 1;9, but that language-specific processes occurred after this age. Major suggested that sounds produced by the child up to the age of 1;9 were actually very similar, regardless of the language from which the word came. This may have had an influence on the conclusions of some investigators that there was an initial single system based on phonetic inventories alone.

Deuchar & Quay (2000) raise similar concerns, and warn that the polarisation of the issue into a question of one *versus* two systems may lead to oversimplification. The authors note that the alternative to two initial systems is not necessarily one initial system. There may be no initial system, especially that there is very little data from the early stages that can be investigated and labeled as belonging to one system or the other. Deuchar & Quay (2000: 113) consider it more fruitful to focus on how and when language differentiation occurs. Their study suggests that it occurs gradually, at different times according to the aspect of language being examined (they examined phonological, lexical, and syntactic development). Lexical differentiation was established early in their subject (around 1;7-1;8), followed by morphosyntactic differentiation (around 1;11) and the emergence of two different voicing contrasts (from 1;11 to 2;3). Within each level of investigation, for instance phonology, Deuchar & Quay showed that the aspect of phonology chosen for investigation may affect whether or not one finds evidence of two systems.

1.3.4 The separate but non-autonomous option

Also recently, some researchers have started wondering whether it is fruitful to persevere with the one-or-two-system issue rather than seeing the two languages of the bilingual as belonging to independent but interactive systems. Bialystok (2001: 103) notes that simple dichotomies, such as whether languages are represented individually or in combination, and whether concepts are linked directly to the second language or mediated by the first, fail to receive empirical support. In adult bilinguals, adequate descriptions of organisation of mental representation include the effect of factors such as level of proficiency and the circumstances of second-language learning. Moreover, it appears that multiple arrangements can coexist in the mind of an individual speaker (e.g. De Groot, 1993). Therefore, there is no reason to expect the situation to be any simpler for children. Efforts to choose one of two possible organisations, for e.g. one system or two, as the defining configuration for children of a specific age (or even specific proficiency level) are doomed to failure. Instead, it is more likely that young children learning two languages experience the same complexity in mental representation as adults do, linking languages and concepts in dynamic ways, and restructuring the systems as needs change and fluency evolves.

Paradis (1998) provided evidence for the separate but non-autonomous option by using an imitation task to compare the truncation patterns of French-English bilingual and monolingual children aged 2;6 on average. The truncation patterns of the bilinguals were similar but not identical to the monolinguals in each language, leading to the conclusion that the bilinguals had separate but non-autonomous phonological systems.

In another study, Johnson & Lancaster (1998) examined the phonological development of a simultaneous English-Norwegian bilingual child between the ages 1;2 and 1;8. Audio recordings were made in different language contexts by different interlocutors (parents or babysitter). Phonetic and phonological analyses consisted of examination of phonetic inventories, prosodic structure, and substitution at the segmental level. Many of the child's early productions had to be discarded from the analysis because they could not be categorised as belonging to only one language, and the authors found it difficult to answer the question of whether the child had developed one or two phonological systems due to the 'similarities' between English and Norwegian. However, they later admit that even cognates such as 'milk' and *melk* are phonetically different and not necessarily perceived as similar by the child, and that acoustic analysis would have helped them better examine some features of vowel quality and length in the child's production in both languages.

Some observations by Johnson & Lancaster included the fact that the child showed preference for English words in an English context and Norwegian words in a Norwegian context, and had translation equivalents, which suggested that he had differentiated the Norwegian and English lexicons. Moreover, the child's phonological development was at the same time different from that of monolinguals his age but also showed similarities with both English and Norwegian monolinguals. For instance, Andreas produced a number of consonants for each language that were found in the production of monolinguals his age, but also other consonants that were not found in the production of monolinguals from either language. With respect to prosodic development, Andreas showed a stronger preference for monosyllables in English than Norwegian and attempted a greater variety of disyllabic structures in Norwegian. His word-final phonetic inventory confirms a more advanced development of coda position in English than Norwegian. The child's phonetic inventory was also larger for English words than Norwegian words, and there was huge variability in his vowel production.

Still, Johnson & Lancaster (1998: 293) noted that there are enough reports of English speaking children to show that there is a wide range of individual differences within this monolingual community and to find a match for Andreas on specific parameters. The same would be true for Norwegian, but the authors point to the need for more studies of monolingual Norwegian development. The authors concluded that the claim that Andreas provided evidence for distinguishing English and Norwegian is different from the claim that he had two separate systems, at least for production. They ask how many levels of phonology should be in place before a system exists, and whether it is necessary to include prosody and features and segments, or whether systemic quality at one level is enough.

The separate-but-non-autonomous option has also been evoked by researchers on successive bilingual acquisition. For instance, Holm & Dodd (1999) followed the development of two successive Cantonese-English bilinguals during their first year of exposure to English (starting at 2;3 and 2;9 respectively). Their phonological process use, phoneme repertoires, and phonetic accuracy were monitored. Both children were found to have separate phonological systems for each language from the start, which, according to the authors, was evident in the following observations: (i) shared phonemes were often used in one language before the other (Cantonese first); (ii) different phonological error patterns were used for each language; (iii) language-specific phonemes were not used in the wrong language; (iv) the same phonemes were differently simplified in each language; and (v) errors always obeyed the phonotactic constraints of the appropriate language.

There was also evidence that phonological development of successive bilingual children is qualitatively different from that of monolingual children. For instance, in terms of phonetic development, both children acquired English voiced plosives before their voiceless counterpart, whereas monolingual English children usually acquire voiceless plosives prior to voiced plosives (Holm & Dodd, 1999: 372). In Cantonese, both children acquired unaspirated plosives before aspirated ones, like monolingual children. Both children acquired all of the other shared aspirated plosives in Cantonese before English. They acquired affricates earlier than monolingual English children, but their acquisition of fricatives was later (Holm & Dodd, 1999: 372). But overall, the phonetic development of successive bilinguals suggests that, because the acquisition of phonemes is due to articulatory maturation, the emergence of sounds is approximately simultaneous in both languages.

The phonological processes by the two children had different profiles. Most of Catherine's processes were shared by both languages, but Max had more language-specific processes. Moreover, both children used phonological processes that are atypical for monolingual speakers of each language, e.g. atypical aspiration and continuant variation of /j, w, l, n/. However, the atypical processes were inconsistent, had only a small impact on intelligibility, and were transient. Moreover, atypical processes were often typical in one language but not the other, e.g. final consonant backing, which is typical in Cantonese but not English, and final consonant deletion, which is typical in English but not in Cantonese, which suggests that the children may have been overgeneralising language-specific rules and applying them in both languages (Holm and Dodd, 1999: 373).

Holm and Dodd (1999: 374) concluded that the types of speech errors and patterns of use that were found in successive bilinguals suggest that the phonological systems of the two languages were interacting. The subjects' acquisition of each language's phonology was qualitatively different from the phonological acquisition of monolingual children from either language. Atypical errors in the children's production only appeared after the introduction of the second language, which suggests that there was an effect on the first phonological system. The authors note that it is possible that atypical errors were caused by an initial inability to process both phonological systems in enough detail to select language-appropriate realisation rules. As both children were exposed to more English, they learned to differentiate the realisational rules for each phonological system. Note, however, that although their Cantonese development was typical before English was introduced, atypical patterns also appeared in Cantonese after the introduction of the new phonological system. Holm & Dodd (1999: 375) suggest that there was an initial negative interference following the introduction of the new system, with overgeneralisation taking place both ways. Perhaps the burden of differentiating each system and abstracting two sets of explicit rules means that for a short period, the established rules of the first phonological system are rethought. Still, the two children kept their two phonological systems separated from the start, but their efficiency in extracting and using the rules of each phonology was initially affected.

1.3.5 Summary

In sum, the bulk of the evidence points in the direction of bilingual children's early differentiation of at least some parameters of their phonological systems. While early studies of phone inventories and substitutions appeared to demonstrate unified language systems, more recent studies have shown that production evidence of the segmental level depends to some degree on articulatory maturation and on a more detailed level of phonetic analysis. Evidence for language-specific voicing contrasts has been shown to emerge by age 1;11 (e.g. Deuchar & Quay, 2000), and segmental contrasts by age 1;8 (e.g. Schnitzer & Krasinski, 1996). Analysis of children's productions at the prosodic level of phonology has yielded evidence that, by about 2 years, children can differentiate production in their two languages by syllable and truncation patterns (e.g. Ingram, 1981/1982; Paradis, 1996; 2000). Finally, a child's phonologies can be differentiated but still show influence from the other language (Johnson & Lancaster, 1998: 271; Schnitzer & Krasinski, 1994).

Despite these results, there has been no clear discussion in the literature with regards to the kind of evidence that is required to establish whether a bilingual child starts with one phonological system or two, or about the nature of the phonological system(s) that the child is expected to acquire. We therefore turn to problems that are intrinsic to the one-or-two-system question.

1.4 Inherent problems with the question

Despite results in recent investigations which are largely positioned towards the notion of each language developing independently from a very early age, some researchers (e.g. Deuchar & Quay, 2000; Johnson & Lancaster, 1998) note that the lack of precise conceptualisation on the nature of 'system', among other issues, make it impossible to determine what type of data would constitute support for separate versus fused systems.

Different researchers have looked at different levels of analysis (phonological, lexical, and syntactic), and, as Deuchar & Quay (2000)'s study suggests, differentiation at each level might become apparent at different ages. Within the phonological level of analysis, different researchers have looked at different phonological aspects in order to answer the question: phoneme repertoires, allophonic distribution, phone trees, phonetic inventories, substitution and simplification, phonological errors, phonological processes, and prosodic features (e.g. syllable-structure, consonant and vowel length).

Schnitzer & Krasinski (1994) tried to specify what kind of phonological evidence one needs before determining whether a child simultaneously acquiring two languages has one phonological system or two at any given time (although in a later publication, they admit that they may have been focusing on segmental repertoires only (Schnitzer & Krasinski 1996: 562)). The authors list three specifications:

> "Spec. A. a unitary phonological system is one in which the child displays any of the following characteristics:
> (i) Failure to use sounds which occur in only one of the two languages
> (ii) Use of sounds impossible in L1 (but found in L2) in L1 lexical items
> (iii) Use of contextual variants (allophones) in the contexts permissible in L1 (but not L2) when using L2 vocabulary.
> Spec. B. On the other hand, in order to claim that two phonological systems have been differentiated, there must be evidence that the child uses the appropriate variants of all phonological classes (i.e. correct allophones of all phonemes) which have thus far been acquired, in all relevant contexts in both languages. Failure to do so in all cases (allowing occasional lapses), would indicate that differentiation is incomplete. The mere correct use of an L1 sound (which did not normally occur in L2) in an L1 word, in itself would not constitute evidence for differentiation.
> Spec. C. In a speaker for whom it has been determined on the basis of Spec. B that the two phonological systems have been differentiated, the use of L1 sounds in L2 and L1 contextual variants in incorrect contexts in L2 must be interpreted as interference. Clearly, it would beg the question to consider whether there is interference between the two systems without having previously established the existence of two systems." (Schnitzer & Krasinski; 1994: 586-587)

Although other researchers have not been as specific about what they mean by one or two phonological systems, many have used similar specifications to the ones listed by Schnitzer & Krasinski (1994). There are many problems with this approach. First, as the authors themselves suggest, the specifications concentrate on segmental aspects only. Some researchers have wondered whether it is appropriate to look at segments at all at an early age, or whether it would be more valid to concentrate on prosodic development. While the early emergence of prosodic features allows investigations of early stages of children's productions, with regards to segmental features, some researchers have argued that language-specific features generally appear late, and that it might be fruitless to try and find evidence for systems at an early age (e.g. De Houwer, 1995; Deuchar & Quay, 1995; Johnson & Lancaster, 1998; Pearson, Navarro & Gathercole, 1995).

These researchers note that, due to constraints on children's articulation in the second year and the difficulty of obtaining sufficient data from young children, it is often difficult to interpret monolingual children's early productions. Determining whether a bilingual child has one phonology or two is complicated by the many crosslinguistic similarities in the composition of early segmental and syllabic inventories and in substitution patterns (e.g. Ingram, 1986; Locke, 1983). Therefore, it is often uncertain whether commonalities between a bilingual's phonologies are due to a unitary system, or due to the lack of language-specific features at that stage in phonological development, which would be apparent in monolingual children as well. These complications highlight the importance of using monolingual controls and of examining phonological properties that show language-specific effects at the age in development being studied (Paradis, 2001: 20).

Second, Schnitzer & Krasinski's first specification concentrates on the potential problem that the bilinguals might face with sounds or realisations of sounds that are part of the inventory of one of their languages but not the other. However, in the case of simultaneous bilinguals, studies on infant speech perception (Section 1.5.1) have actually shown a remarkable ability to gradually tune into the sounds of the ambient language(s) in the first year of life, and to build up phonological representations accordingly. This undermines any suggestion that the bilingual might fail to use sounds which occur in only one of the two languages, as the infant's perceptual abilities, which constitute a prerequisite for its production abilities, are not initially set for one language or two. As for successive bilinguals or second language learners, recent research shows that they might actually have more problems with sounds that are 'similar' in the two languages (e.g. Flege, 1995) than sounds that are exclusive to one of the two languages. However, this approach is not without problems, due to the difficulty of establishing a basis for crosslinguistic phonetic similarities (see discussion in Strange, 1995; 1999).

Third, with regards to specification B, a contrastive analysis of the phoneme inventories of the two languages and of their allophonic distribution does not contain enough detail about the articulatory or acoustic structure of phonetic segments in each language to allow one to make informed predictions about whether the child has acquired the sound systems of the two languages or about possible difficulties they might encounter. While in simple terms we might talk about the 'system' of English and the 'system' of Arabic, it is clear that each system is only identifiable in a rather general abstract sense. For example, a phoneme like /t/ might be judged as the 'same' in two languages, despite important phonetic and/or phonological differences that may govern its production in each language and differences in the system of oppositions and functional load. These include place and manner of articulation, phonotactic distribution, systematic social and stylistic differences, and subtle differences in articulatory coordination that may not be detected through auditory analysis alone.

If we focus on an aspect of the phonological system of English such as /t/, it is hard to define exactly what evidence we need to look for to decide whether a child has successfully acquired it. For example, in English, /t/ varies in its phonetic realisation according to word-position. Moreover, it varies systematically across dialects and even within dialects. The notion of the phoneme itself has been questioned, since there is no default realisation for something like a /t/ which all speakers are equally likely to produce. For instance, in their study of /t/ production in Tyneside English among adults from the same local community, Docherty, Foulkes, Milroy, Milroy, & Walshaw (1997: 293) found a strong correlation between patterns of production of five different variants of /t/ ([ɹ], [t̬], [t], [ʔt], and [ʔ]) in word-final pre-vowel position ('get it', 'lot of'), and social factors such as age, sex, and social class. The authors concluded that if social dimensions are not taken into consideration, then 'an account of the complex alternations of word-final pre-vowel (t) in Tyneside English based on data from a group of middle-class men would be likely to draw very different conclusions from one whose observations were based on older working-class women' (Docherty & Foulkes, 2000: 111). In assessing how something like /t/ is acquired by children in that community, one therefore has to take such facts into account in order to define targets accurately. A follow-up study examining to what extent these detailed accentual features were being acquired by young Tyneside children aged 2-4 (Foulkes, Docherty and Watt, 1999) will be discussed in Section 1.5.3.3.

The surge in cross-linguistic studies of language acquisition and, in some cases, the use of advanced instrumental analysis techniques, has also shown that an abstract phonemic approach does not capture important language-specific allophonic (e.g. Ball, Muller, & Munro, 2001; Deuchar & Clark, 1995; Hazan & Boulakia, 1993; Holm & Dodd, 1999; Watson, 1995), phonotactic (e.g. Johnson & Lancaster, 1998), and prosodic (e.g. Grabe, Post & Watson, 1999; Paradis, 2001; Vihman & Velleman, 2000; Whitworth, forthcoming) patterns of variation.

The introduction of instrumental techniques in the study of bilingual and second language speech has shown that detailed language-specific phonetic features that are involved in the production and perception of sounds undermine the phonetic 'similarity' often assumed by phonemic analyses (e.g. Flege, 1995; Strange, 1999). Deuchar & Quay (2000: 29) draw attention to the fact that a great deal of phonological analysis of child speech depends on transcriptions and phonetic judgements by the analysts. Although extensive training and reliability checks help to reduce possible errors, it is sometimes useful to make acoustic as well as perceptual analysis. Similarly, Pearson & Navarro (1996) point out that acoustic studies have been particularly useful in identifying language-specific differences in early bilinguals, given that a segmental approach is limited by the late acquisition of language-specific differences. Acoustic studies using instrumentation have the advantage of providing information that is not always perceptible to the average listener or even by a trained phonetician.

Another problem concerns the overwhelming emphasis on the acquisition of contrasts in bilingual (and monolingual) phonological acquisition when determining whether a child has successfully acquired the phonological system(s) of his/her language(s). For instance, Deuchar & Quay's (2000) criticism of looking at phonetic inventories is that they 'do not reveal much about the nature of the system in terms of contrast and oppositions'. Similarly, Johnson & Lancaster (1998: 271) point out that if we recognise children's own way of marking contrast, for example with subphonemic VOT distinctions, language-specific substitutions for phonological segments that occur in both languages, word-truncation patterns, we are likely to see differentiation. These statements emphasize the primacy of phonological contrasts in the approach to language differentiation by the child, which constitutes only one aspect of the nature of child phonological acquisition. Insights from recent variationist sociolinguistic studies of monolingual acquisition have shown that different types of variability in the speech input that a child is exposed to such as dialectal, individual, and stylistic differences constitute part of the knowledge acquired by children (these will be discussed in Section 1.5.3).

With regards to Schnitzer & Krasinski's third specification, the term 'interference' is problematic when evoked out of the social context in which the so-called interference took place in the child's production. This issue deserves more attention if one is interested in a better and fairer interpretation of bilingual speech behaviour, and will be discussed in more detail in Section 1.7. Moreover, with respect to very young bilinguals still undergoing the process of language acquisition, it makes little sense to talk about interference at all, since neither of the two systems is fully established yet (Hoffmann, 1991: 95).

Finally, as Johnson & Lancaster (1998) point out, before asking whether a developing bilingual has one or two phonological systems, there are a few prior questions that apply to the study of monolingual phonology. Such questions include whether we are talking about comprehension or production, when any child can be considered to have a phonological system, and what we are assuming about the nature of the phonological system in the lexicon.

It is difficult to define a system even in monolingual acquisition, due to the debate over what a phonological system looks like and what age it emerges at (Burton-Roberts, Carr & Docherty, 2000). In broad terms, a phonological system represents the speakers' knowledge of the sounds of their native language, along with the features that enable them to produce and comprehend the systematic patterns of that language. However, the nature of these sounds and of their systematic patterns is a matter of debate in phonological theories due to dispute over the relation between their underlying phonological representation, which is generally argued to be invariant, and the variable phonetic output (Docherty, 1992: 56). The idea of a unique phonological representation for any linguistic output is also being challenged in recent work on perception and production on the one hand, and sociolinguistic studies with a variationist perspective on the other. Each of these issues will now be discussed in detail in Section 1.5, along with their implications to the study of bilingual phonological acquisition.

1.5 Issues in monolingual development

This section aims to discuss controversial issues in monolingual acquisition which may have added to the divisive views on the nature of bilingual development. These issues relate to the development of perceptual (Section 1.5.1) and productive (Section 1.5.2) abilities in monolingual children, along with the role of linguistic (Section 1.5.3) and social (Section 1.5.4) variability in shaping the nature of the phonological representation in the monolingual child. The discussion is by no means intended to be comprehensive in its coverage of all the stages of phonological development, but rather concentrates on issues that will be evoked in this study for a better interpretation of bilingual phonological behaviour.

1.5.1 Early development

According to current theories of L1 phonological development (e.g. Best, 1995; Jusczyk, 1993; Kuhl & Iverson, 1995; Werker, 1995), infants' perception of speech begins to shift in the first year of life from a language-universal pattern to a language-specific pattern of organisation which reflects the phonological structure of the ambient language. Cross-language studies of discrimination of native and non-native contrasts by infants suggest that infants are 'universal perceivers' (Strange, 1995: 19), i.e. phonetic contrasts are perceptually differentiated, regardless of their phonological status or even their occurrence in the adult language to which infants have been exposed. At this early age, perception is not yet affected by specific linguistic experiences, but rather, reflects young infants' predisposition to detect specific, maximally-contrasting, temporal patterning as well as distributional regularities in the input (Saffran, Aslin, & Newport, 1996; Vihman, 1996). Between early infancy and adulthood children's interactions with their linguistic environment produce significant changes in the perception of speech sounds (Section 1.5.1.1). This has important implications for bilingual language acquisition, since it suggests that the neural mechanisms underlying human language acquisition are not necessarily initially set for one or two languages (1.5.1.2).

1.5.1.1 Early perceptual abilities

From the moment children are born, they show sensitivity to speech sound recognition. Moon, Cooper, & Fifeer (1993) observed a preference in newborns (two days old) for listening to their native language. Similarly, Mehler, Juczyk, Lambert, Halsted, & Betoncini (1988) showed that two-month-old infants discriminate between two unknown languages that belong to relatively distinct language families, while Bosch & Sebastián-Gallés (1997) showed that four-months-old infants also distinguish between two closely related families (Catalan and Spanish). Moreover, there is evidence that at this early stage, infants are able to discriminate contrasts that do not appear in the language spoken in their native environment (e.g. Aslin & Pisoni, 1980; Eimas, Miller & Jusczyk, 1987). Infants' capacities for perceiving speech also go beyond discriminating one kind of syllable from another. They are able to compensate for differences in talkers' voices (Kuhl, 1993), they appear to recognise their own mother's voice, and they even seem to know when speech is being directed to them rather than to an adult (e.g. Cooper & Aslin, 1990).

Within the first six months of life, infants are prepared to accommodate to any language-particular selection from the universal set of possible phonetic categories. In the latter half of the first year, however, the location of the 'natural' phonetic boundaries may undergo shifts as a function of specific linguistic experience. Werker, Gilbert, Humphrey & Tees (1981) have established the timing of the shift from broad discriminatory abilities to more adult-like language-particular biases as late as in the first year for consonantal contrasts, while recent work by Kuhl & Iverson (1995) & Werker (1995) suggest an earlier change in orientation for the more salient vowel categories. Global properties of speech such as stress patterns, syntactic juncture, and intonational contours are attended to even earlier, as Jusczyk, Hohne & Mandel (1995) have found that infants recognise these patterns in the very early months of life, and later begin to attend to the fine-grained structure of native-language phonetic sequences.

Jusczyk et al (1995: 114) found that infants make some important discoveries about the organisation of native language sound properties between four and a half and nine months of age. The growth of knowledge regarding the native language occurs at the same time as the apparent decline in sensitivity to certain foreign language contrasts. The infants learn to attend to those aspects of the speech signal that are critical for distinguishing among words in the native language. Perceptual dimensions that are attended to are 'stretched', allowing infants to make finer distinctions, whereas unattended dimensions are 'shrunk', making them harder to perceive (Kuhl & Iverson, 1995). Infants are not only sensitive to the more global aspects of the sound structure of the native language, but are learning a great deal about the fine-grained features of the sound structure of the native language (Jusczyk et al, 1995: 114).

By the end of the first year of life, infants' phonetic perceptual sensitivities reflect considerable influence from the native language (Jusczyk, 1993). This influence is evident both from a preference for highly frequent phonetic patterns, and in narrowing of initial discriminatory abilities to match the contextual distribution of phonetic information in the input. Thus, native language patterns are well-established long before children have mastered the production of the phonetic segments and sequences of that languages.

The mechanism for the shift towards the phonological patterning of the native language remains controversial. Vihman (1996: 96) suggests that the role of development in motoric and 'motivational' systems (that is the development of vocal production and intentional communication) offers some answers. The role of the mother's voice, then of the adapted intonation patterns instinctively used in addressing infants, may guide the infant towards specifically language-relevant syntactic units of the native language.

Kuhl (1993) describes the infant as a 'citizen of the world' and the adult as 'culture bound' due to the fact that, as we get older, our abilities to differentiate the sounds of the world's languages are greatly reduced. Kuhl uses the 'perceptual magnet effect' to explain how adults' and infants' phonetic perception is altered as a function of exposure to language. Her Native Language Magnet (NLM) model argues that exposure to language early in life produces a change in perceived distances in the acoustic space underlying phonetic distinctions, and this subsequently alters both perception of spoken language and its production (Kuhl, 1994). According to NLM, older infants' and adults' internal representations of phonetic categories reflect a language-specific 'warping' of the multidimensional acoustic-phonetic space such that within-native-category acoustic differences are perceptually shrunk around category prototypes, while between-native-category acoustic variations are perceptually stretched at phonetic boundaries. This will cause certain perceptual distinctions to be maximised (those near the boundaries between the two magnets) and others to be minimised (those near the magnet attractors themselves).

1.5.1.2 Implications for bilingual acquisition and L2 learning

With respect to the infants' initial speech discrimination abilities, similar abilities have been found in bilingual infants. For instance, research on speech perception in children raised bilingually (Catalan & Spanish) indicates that they can discriminate different language-specific phonological contrasts as early as four and a half months of age (Bosch & Sebastián-Gallés, 2002). Moreover, following analysis of the perception of phonemes by four-to-eight-month-old infants raised in bilingual (Spanish-English) or monolingual

(English) environments, Eilers, Gavin & Oller (1981) found that the former discriminate better than the latter not only between English and Spanish phonemes, but also between the phonemes of English and those of Czech, a language to which they had never been exposed. The authors interpret these results as possible evidence that a richer linguistic input from the environment fosters a better development of the relevant skills, in this case phonemic discrimination.

Therefore children exposed to two languages appear to perceptually discriminate different linguistic systems at birth, a capacity that is a prerequisite to establishing different representations of two languages. The ability of infants to gradually tune into the sound patterns of their ambient language(s) from the first year of development suggests that there are no grounds for the possibility that simultaneous bilingual children will have problems perceiving sounds that are part of only one of the two languages. Moreover, there is no basis for the assumption that they will perceive sounds from the two languages as 'similar' since the studies reviewed above suggest that infants attend to detailed phonetic patterns in the input that they receive.

As for successive bilinguals, if there is a loss of discrimination ability for non-native contrasts between infancy and adulthood (Werker et al, 1981), the puzzling question is how the acquisition of a second language in early childhood appears to result in native phonological fluency (e.g. Flege, 1995). The contrasts which have supposedly been filtered out are nevertheless acquired by bilingual second language acquirers. In recent research, Werker (1995) notes that their earlier conclusions was not accurate. A decline in phonetic perception by one-year olds is not attributable to a general decline in 'auditory attention', but rather reflects the development of selective patterns of operation. It also appears that the perceptual difficulties of non-native listeners do not result from a loss in the sensory capacity to detect acoustic differences that are not used in contrasting phonemes in the native language. Werker & Tees (1984) and Werker & Logan (1985) later showed that adult listeners can discriminate even the most difficult non-native contrasts with much the same accuracy as native listeners. Thus the ability to detect phonetically relevant acoustic variations in speech utterances is not irretrievably lost in the course of learning the native language.

With regards to Kuhl's NLM model, work on adults suggests that the boundaries between the magnets do not literally disappear; it is possible to increase performance on the discrimination of foreign language contrasts in adults through extensive training (Flege, 1995). For instance, perceptual studies of adult L2 learners provide encouraging evidence that, at any age, modification of phonetic perceptual patterns is possible. Second language learners with extensive immersion experience or intensive conversational training show marked improvement in ability to differentiate perceptually even the most difficult non-native phonetic contrasts (e.g. Pisoni & Lively, 1995).

1.5.2 Early production abilities

1.5.2.1 The prelinguistic period and the relation of babble to speech

Research on prelinguistic vocal development has two recurring themes. First, regardless of the language community in which they are raised, infants pass through an ordered sequence of stages in terms of vocal development. The stages of vocal production differ from one model to another and are difficult to delineate due to individual differences, but the emergence of canonical syllables (or 'templates') is common to all models (Vihman, 1996: 118-120). Second, due to the similarity of infants' first vocalisations (Menn & Stoel-Gammon, 1994: 338), it is difficult to determine what vocalisation belongs to what system or language, if any. The debate on the 'babbling drift' shows that it is by no means clear whether even monolingual infants show target language effects in their babbling or not (e.g. de Boysson-Bardies & Vihman, 1991).

Initially, infant vocalisations are impulsive and unstructured productions; by the second month, some comfort state 'coos' and 'goos' emerge. At the onset of canonical babbling, infants' utterances become increasingly adultlike with identifiable CV syllables and clear intonation patterns (Menn & Stoel-Gammon, 1994: 338). The role of the social context in facilitating development in vocal production is still unresolved. In some accounts, the child is considered to be motivated by the need to exercise abilities and play rather than any conscious effort to learn to talk (e.g. Stark, 1980: 90). Others see the transition from the first to the second stage of vocal production as evidence for the relationship between early social interaction and infant vocalisation (Vihman, 1996: 118-120).

With respect to babbling, the sounds and syllable structure characteristics of the later babbling period (10-12 months) are highly similar across subjects and across languages. For example, in all studies so far, the consonantal repertoires of infants in the later babbling period typically include a high proportion of front (labial and dental/alveolar) consonant, of stops and nasals, and of CV syllables (Menn & Stoel-Gammon, 1994: 338). Due to the similarity of infants' first vocalisations, it is difficult to determine what vocalisation belongs to what system or language, if any.

Counter to Jakobson's (1968) discontinuity claim that babbling and phonemic development at the onset of speech are unrelated, more recent research indicates that there may be a drift in the structure of babbling towards the sound patterns of the ambient language(s). Researchers like Locke (1983) have shown that the sounds and syllable structure characteristic of the canonical babbling period closely resemble those of early meaningful speech. This development may start during the second half of the first year of life (e.g. de Boysson-Bardies & Vihman, 1991), and is due to the fact that infants are capable of creating mental representations of sound categories without reliance on either lexical items or knowledge of abstract phonemic principles before the onset of speech. Changes in babbling behaviour become particularly noticeable around 8-10 months, including an increase in the number of sounds which also occur in the target language (e.g. Vihman, Ferguson, & Elbert, 1986; de Boysson-Bardies & Vihman, 1991).

Individual differences during the babbling period in terms of place and manner of articulation, syllable shape and vocalisation length provide further support for the continuity between babbling and speech; these differences are often 'carried forward' from the prelinguistic period to the first words (Stoel-Gammon & Cooper, 1984; Vihman, Ferguson, & Elbert, 1987). For instance, Vihman (1992) reports individual differences in the occurrence of 'practiced syllables' in babbling and then shows that these same syllables form the foundation of children's first words. Moreover, crosslinguistic research by de Boysson-Bardies, Vihman, Roug-Hellichius, Durand, Landberg, & Arao (1992) shows language-specific effects in both the consonantal and vocalic system of the ambient language in the prelinguistic vocalisation of infants as early as ten months.

Babbling and practice provide the infant with feedback from their own vocal input and from caretakers. By listening to their own productions, children establish a link between their own oral-motor gestures and the acoustic signal which results; this is known as 'feedback loop' and is a prerequisite to auditory-vocal matching which underlies word production (Menn & Stoel-Gammon, 1994: 339; Vihman, 1996: 119). The feedback loop may help children recognise words in the adult language that resemble their babbled forms, e.g. 'ball' for [ba] or [baba]. But visual as well as auditory factors enter into the child's first expression, in production of features of the ambient language. Research has shown that infants pay attention to the visual effect produced by talking faces (e.g. Meltzoff & Moore, 1983). For the infant these visual cues are taken from the caretaker from the earliest moments of social interaction.

With respect to prosody, there are elements which appear to be naturally available to the infant in the prelinguistic period, e.g. voluntary modulation of pitch and final syllable lengthening, leading to the beginnings of an intonational system which appear late in the prelinguistic period but which begin to coherently develop only with the first steps in syntactic structure (beyond the one-word stage). The control of pitch increases and stabilises throughout the first year of life. The predominant prosodic characteristics of the adult system are reflected in infant productions within the one-word period at the latest, when only a subset of the adult segmental inventory may be used; acquisition of the full system is not typically achieved until after the child has begun to master the syntactic system, however (Vihman, 1996: 212).

1.5.2.2 Early units of production

Children's earliest phonological 'units' appear to be whole words (e.g. Ferguson & Farewell, 1975; Vihman, 1994; Wode, 1997), although other authors have treated the syllable (e.g. Moskowitz, 1973) or isolated phonemes (e.g. Jakobson, 1968) as subword units of construction (see discussion in Vihman, 1996).

First words are identified when the child begins to produce existing phonetic patterns developed through babbling in situations appropriate to similar (or matching) adult word patterns. The first words of early talkers may not appear to be phonologically related. Each is the product of an idiosyncratic match between a prelinguistic 'gestural score' or 'articulatory routine' and a salient adult word (Vihman & McCune, 1994). Early in the second year, the child experiences an expanded capacity for internal representation; it is hypothesised that this maturational change provides the necessary basis for phonological systematisation, in which one or more word templates are formed and used to assimilate growing numbers of adult forms (Vihman & McCune, 1994).

It is in this second stage of phonological organisation that the child begins to accommodate adult forms which go beyond his or her production constraints by making systematic changes in the reproduction of adult segments, sequences and syllable shapes. These adjustments have been termed (child) phonological rules (Menn, 1971; Smith, 1973), or processes (Stampe, 1979; Oller, 1975), despite the fact that the relationship to rules of adult phonology may be more apparent than real. Traditional descriptions of these rules and processes have emphasised their universality and assumed that they encompass the entire lexicon (e.g. Jakobson, 1968) or operate across the board (Smith, 1973). However, recent attention in crosslinguistic studies to individual differences in children's productions and in the input they receive have undermined the universality of these processes.

For instance, Ingram (1979) discussed the limitations of the concept of phonological processes by looking at phonological patterns in the speech of young children aged 1;6 and 4;0 and from different language backgrounds. The main limitation concerned accounting for individual variation from one child to another and the role of the phonological system of the language that the child is acquiring in the application of these rules or processes. For example, gliding is more frequently documented in English than in French, while denasalisation is characteristic of French. Similarly, Vihman (1978), found consonant harmony to range in the use of thirteen children, from 1% for a Chinese-speaking child to 32% for an English-speaking child. Vihman notes that consonant harmony may therefore play a negligible role in some children's phonological development, although its use is the best documented and most discussed phenomenon of early child phonology (Menn, 1971; Cruttenden, 1978; Vihman, 1996). Wode (1997) adds that there is no child who harmonises across-the-board; some children do not even harmonise all tokens of a given word during the same recording session in the same way.

Wode (1997) analysed production data from German infants aged 0;7-2;3 and found no support for the common view that early child phonology is based on phonological processes. The main finding from the data was the great amount of variation in early L1 production. When first acquired, the pronunciation of a target lexical item may be close to the target, but there are exceptions. There is likely to be variation of different sorts, including substitution patterns, which tend to be anything but identical across different children (Wode, 1997: 21). Wode offers three explanations for variation in children's productions: (i) lack of motor control, (ii) processing, and (iii) perception.

With respect to lack of motor control, a considerable portion of early child phonological variation is very likely due to insufficient motor control. This can be assessed in terms of gestural phonology (e.g. Browman & Goldstein, 1992), which stresses the importance of articulatory gestures for the description of speech and phonological analysis. Thus, insufficient stop gestures produce continuants; close misses of the place of articulation, such as alveolar, may result in dentalised variants; variation between voiced and voiceless derives from the lack of phonation control; or variation among aspirated versus unaspirated phones results from insufficient control of aspiration (Wode, 1997: 25).

With respect to processing, phenomena like harmony of places of articulation (e.g. doggie [gɔgɪ]) can be explained in terms of anticipation of the place of articulation from some subsequent position in the target or retraction from a preceding one. Such errors are labelled 'dislocations' or 'deviations'. Wode argues that some deviations may be due to the way mental representations are created in memory and activated for production. It is well known that no two tokens of a word are acoustically identical. If a range of variation exists in carefully elicited speech (e.g. Peterson & Barney's 1952 vowel data), then the range of variation in real-life speech as input to children should at least be the same, if not much larger. Recent research (Docherty et al, 2002) has actually shown that variability in child production was linked to the mothers' production variability in child directed speech. This issue will be discussed further in Section 1.5.2.1.

With respect to perception, Wode (1997) uses findings from the abilities of infants to tune into the sounds of the ambient language in the first year to monitor and/or control the development of production. Wode suggests that the speaker's mental representations of the target language(s) play a crucial role in monitoring production. With respect to variation in L1 production, certain kinds of variation may be a direct effect of how infants/children perceive their input and create mental representations in memory. If the input presented to children is as varied as suggested by Docherty et al's study, the question is how L1-learning children are to determine the appropriate phonological representation of a given lexical item. Wode (1997: 36) suggests that children can only take the actual token at face value and store it according to their categorisation abilities. Subsequent tokens are then stored in such a way as they are superimposed on the previous ones. All instances of a given lexical item are further organised into a network and, as time goes on, the less central representations will be outnumbered by the more central ones. This process may lead to 'fuzzy' representations in the sense that lexical representations do not necessarily have clear cut boundaries. In activating the representation of a given lexical item from production, a chid may not always hit upon the central part of the representation so that less central parts may surface in the shape of the child's substitutions as described in the acquisition literature in terms of phonological processes or equivalent terms (e.g. Ingram, 1979; 1989). Wode's assumption is that they are not processes at all in the sense that children change anything; they simply activate different parts of their fuzzy representation.

One important consideration in studying a young child's speech is that the system observed in under constant change, showing older and newer developments at any time. This is highly important in the study of phonological development, and is manifested in a variety of ways. One of the most striking consequences is the phonetic variability that children show in their pronunciation of words. Children will often show a variety of productions for the same words. One reason for phonetic variability is presumably the fact that children are gradually moving from one pronunciation of a word to another. The more complex a word is, and the more recent has been its acquisition, both appear to contribute to higher variability in pronunciation. At the same time, there is a simultaneous occurrence of advanced and frozen forms, which shows the dynamic nature of the child's system, but also makes it difficult to generalise about phonological processes since one must consider the words which they affect.

Ingram (1979) suggested a distinction between the adult's pronounced form, the child's perceived form, the child's underlying form, and the child's spoken form. Children therefore have representations for both the adult form and for their own form, and the latter might become resistant to phonological processes. The child might also show a phonological preference for a particular class of sounds, such as fricatives or nasals, or a particular kind of syllable structure. As a result the child will produce an unusual number of words that show the preferred sound or syllable structure. Preferences like these can lead to individual variation. Phonological development in this stage therefore consists of both general processes and also of unique phonological preferences that children show in various productions of speech forms for the language they are acquiring (Ingram, 1979: 138-148).

1.5.2.3 From words to segments

At a later stage in the child's development, there is a gradual qualitative shift from a predominance of processes affecting the structure of whole words (consonant harmony, reduplication, final consonant deletion) to those affecting specific segments or classes of segments (stopping of fricative, gliding of liquids). The shift itself can be understood as an indication of the gradual emergence of segments as control units for the child (Vihman, 1996: 216).

Studdert-Kennedy (1987: 67) argues that the shift from the word to the segment is the result of vocabulary growth, which leads 'recurrent patterns of sound and gesture to crystalise into encapsulated phonemic control units', resulting in emergence of a full repertoire of phonemes by the middle of the third year. Similarly, Nittrouer, Studdert-Kennedy, & McGowan (1989) note that, as the number and variety of words in a child's lexicon increase, words with similar acoustic and articulatory patterns begin to cluster; from these clusters, coherent units of sound and gesture (or phonetic segments) eventually emerge. Like Studdert-Kennedy (1987), Nittrouer et al (1989) note that the emergence of segments is a gradual process, perhaps beginning as early as the second or third year of life when the child's lexicon has more than 50-100 words. But the process is evidently still going on at least in some regions of the child's lexicon and phonological system as late as seven years of age (Nittrouer et al, 1989: 131).

Lindblom's (1992) functional model suggests that segments of a later stage of the child's phonological organisation emerge through the 'interaction of subsystems' in the form of a build up of distinct word forms (or gestural scores) involving different activity patterns for the various articulators. Because the structure resulting from this interaction is self-organising, 'children are never aware of having acquired phonemic coding. It appears to emerge in a completely automatic and implicit manner' (Lindblom, MacNeilage, & Studdert-Kennedy, 1984: 185).

The age and the order of acquisition of phones, phonemes, and phonemic contrasts is variable across children and only probabilistic statements can be made. There is no typical or universal order of acquisition for children learning a given language. Some phones (contextually-determined variant pronunciations) are acquired earlier than others. Dialect differences have generally been ignored, as have other types of variation (Menn & Stoel-Gammon, 1994: 347). Furthermore, longitudinal studies of pronunciation have found considerable individual differences in the order of acquisition of sounds such as stops (Macken, 1980) and fricatives (Edwards, 1978). The order and age of mastery of phonemic contrasts is likewise variable across children within a given language. Many exceptions to the order of contrast acquisition which was proposed by Jakobson have been found in the literature. Therefore, his famous laws of 'irreversible solidarity' cannot be considered tenable.

1.5.3 Later abilities

Most of the research into children's phonological development has concentrated upon the first five years of life. Moreover, most investigations have been concerned with establishing the patterns in the development of segmental phonology. It is generally asserted that by age 5;0, the majority of children have developed effective abilities in their use of spoken language (Ingram, 1976: 44). However, phonological development arguably continues throughout later childhood (Grunwell, 1986).

With respect to perception, even though children demonstrate discrimination abilities from an early age, their mastery of perceptual distinctions is not fully mature. Fourcin (1978) reports experiments which indicate that the establishment of phonemic categorisation skill continues well beyond 5;0 and that it may be as late as 14;0 before children begin reliably to display sharp categorical responses to certain synthetic acoustic stimuli simulating distinctive features of the speech signal. As for production, the acceptable pronunciation of certain English consonants is not achieved until between about 4;6 and 6;0 (Ingram, 1976). The sounds most commonly listed as latest to master are /θ ð ʒ/ followed by /ɹ z v/ and affricates (e.g. Sander, 1972). There is agreement amongst the results of most studies that children complete their phonemic inventory by the age of 6;0, or at the latest 7;0, with the mastery of the pronunciation of these last few consonants (Grunwell, 1986: 36). Certain segments continue to present them with articulatory problems, particularly the fricatives /s ʃ/, which are characteristically palatalised in children's speech even after 5;0. Articulation of consonant clusters is another aspect of pronunciation which some children take considerable time to master. Vihman (1996: 237) notes that the fully mature segmental organisation is not complete until well into the 'grammar school' years, presumably around age 11.

Apart from pronunciation maturation, children gradually come to know which phonetic segments are phonologically distinctive, which phonetic variants are appropriate in which contexts (allophonic constraints), which phonetic sequences are 'allowed' (phonotatic constraints), and how phonetic segments vary in different lexical and sentence contexts (prosodic constraints) and in different styles of speaking (e.g. careful *versus* casual speech constraints) (Lindblom, 1990).

With respect to prosodic development, Vihman (1996: 235) notes that it is an important element in the transition to syntax. For instance, prosody has an important role in the development of morphosyntactic structure. Elements often omitted in child speech (e.g. articles) lack prosodic salience, or stress in English (Brown, 1977). The initial tendency to omit function words may be additionally guided by the lexical template developed earlier (within the one-word stage), as a response to the greater prosodic salience of final syllables or the dominant pattern of early content words addressed to the child (or both). Such a template would then gradually fade as the child acquires the specific rhythmic structures of the target language as well as greater knowledge of and facility with morphosyntactic structure (Vihman, 1996: 235).

Very little is known about how prosody develops in later childhood. One probable reason is the lack of an agreed framework of analysis. Another reason is the difficulty of defining the 'meaning' of prosodic contrasts, which often signal the more indefinable aspects of communication, such as attitudes and emotions (Grunwell, 1986: 42). There are, however, discrete grammatical functions signalled by prosodic contrasts and it is these which have been investigated in the few studies of children's prosodic development that have been conducted. Prosodic development continues throughout childhood (Crystal, 1986). For instance, the nearer a child is to 12;0, the more likely he or she is to have control of his/her stress placement rule (e.g. *ˈgreenhouse*, *green ˈhouse*). The establishment of this control is gradual, with considerable individual variation as the rule is induced. Intonational contrasts are a bit more difficult to acquire and comprehend, and children between the ages of 7;0 and 10;0 are still in the process of acquiring the fundamental functions of English intonation, especially for signalling grammatical contrast and for taking account of the situational context (Grunwell, 1986: 45).

So far the discussion has mainly concentrated on children's output rather than on the nature of the input that they receive, and on their abilities to acquire phonological contrasts with no attention being given to other aspects of phonological knowledge. Until recently, the task of the child was seen as being to acquire the full inventory of adult oppositions between contrasting sounds (Ferguson, 1976: 84). For linguists viewing phonological development as the acquisition of phonemic oppositions (Jakobson, 1968), or the refinement of realisation rules (Smith, 1973), variability was an inconvenience to be acknowledged but not attended to. Similarly, child phonologists were not interested in the acquisition of the range of phonological variation that marks the speaker as coming from a particular region or social groups, or that marks a conventional style or register of the speech community, although these are clearly aspects of language use that need to be learned. Variability in the input is present from the beginning, reflecting in part dialect and register differences, but the reaction of the child to such variation has, until recently, only occasionally been noted (Local, 1983). The quote from MacKain & Stern (1985) below captures what many researchers consider to be the essence of phonological development.

> "Essential to language development is the discovery of those sounds that contrast in the target language to convey differences in meaning. In acquiring these oppositions, the language user establishes phonemic categories. The speaker is perceptually sensitive to the acoustic parameters that function to distinguish these categories while remaining relatively insensitive to the parameters that do not distinguish meaning... the infant must eventually come to recognise and construct an internal representation of phonetic oppositions with phonemic significance and also assimilate phonologically irrelevant phonetic variations to represent a single phoneme." (MacKain & Stern, 1985: 1-3)

There are two problems with this approach. First, it assumes that phonetic variation in the speech output is irrelevant because it does not contribute to meaning (or even provides a barrier to clear conveyance of meaning), and therefore it is not part of the assumed invariant underlying representation. Second, it assumes that the child's job is to acquire the phonemic oppositions that are relevant to the construction of meaning in its environment. Both these assumptions have been challenged in work on variability in speech perception (Section 1.5.4) and on sociolinguistic variability (Section 1.5.5).

1.5.4 Variability in speech perception

1.5.4.1 Types of variability in the input

Normal hearing listeners can adapt easily to changes in speaker, dialects, speaking rate, and speaking style, as well as a wide variety of acoustic transformations, including phonetic context, prosodic contours, and the presence of noise (Pisoni, 1997: 16; Goldinger, 1997: 34). Pisoni (1997) and Perkell & Klatt (1986) list the following types and sources of variability that is available in speech:

- Ambient conditions (e.g. background noise, room reverberation)
- Within-speaker variability (e.g. breathy/creaky voice quality, shifting formants, changing speaking rates, imperfect repetitions across tokens of the same gesture)
- Cross-speaker variability (e.g. differences of dialect, vocal tract length and shapes, detailed articulatory habits)
- Segment realisation variability (e.g. coarticulatory changes, articulatory modification due to stress or duration changes)

- Word environment variability in continuous speech (e.g. cross-word-boundary coarticulation, phonetic and phonological recoding of words in sentences, changes in word duration due to syntax, pragmatics).

Two other types of variability can also be added (Hirschfeld & Gelman, 1997: 214):

- A rich array of differences in register associated with the speaker's social group (e.g. gender or occupational), social role (e.g. subordinate), or social context (e.g. classroom vs. playground)
- Difference in the language itself (e.g. English vs. Spanish).

Pisoni (1997) studied the effect of stimulus variability from different talkers and different speaking rates on word-recognition performance. Recent findings have suggested to him that some of the early theoretical assumptions that speech researchers have held about the existence of abstract units such as phonemes and words need to be reexamined and substantially revised. More specifically, researchers have looked for physical invariants in spoken language in their search for underlying representations and tended to ignore the problem of stimulus variability in the listener's environment. Variability was simply treated as a source of 'noise' in the acoustic signal (Pisoni, 1997: 10).

Traditional accounts of speech production and perception (e.g. Chomsky, 1965; Jakobson & Halle, 1956) emphasise that canonical linguistic representations are derived from the speech signal. In these accounts, after the system makes a response to variation in the signal and derives a canonical representation, information about nonlinguistic variation is discarded. This view implies that each item in the mental lexicon consists of only one phonetic form. Johnson & Mullennix (1997) call this the 'mental dictionary assumption' because, in this view, words in the head are exactly analogous to dictionary entries. The mental dictionary makes speaker normalisation necessary due to the large acoustic differences between speakers; if each word is stored in the mental lexicon with only one canonical phonetic form, then the main problem in speech perception is to take speech signals that do not exactly match the canonical form and transform them so that sources of variation (such as vocal tract length and vocal fold vibration pattern) are eliminated (Johnson & Mullennix, 1997: 2).

In addition to talker variability, speech displays a large amount of contextual variability. In traditional accounts, however, each speech sound is considered to have a unique context-independent feature description, like a dictionary entry. Johnson & Mullennix (1997: 3) call this alphabetic writing assumption because we write the initial *d* in *Dee* using the same letter as the initial *d* in *do* and they are assumed to be merely positional variants of the same linguistic unit. This contextual variability problem, typically described as the 'invariance problem', has occupied the attention of speech researchers and prompted the development of a wide variety of theoretical viewpoints (see Johnson & Mullennix, 1997 for a review). Johnson & Mullennix (1997) suggest that a rich-representation, simple-mapping approach to talker variability (rather than the traditional simple-representation, complex mapping approach) might be fruitful in dealing with contextual variation. This issue will be explained below.

Research on auditory word recognition (Pisoni, 1997: 10) suggests that stimulus variability is 'lawful' and 'informative' for perceptual analysis. Pisoni reviewed some recent experiments on talker variability and perceptual learning that indicate that listeners encode fine stimulus details about the talker and use them during word recognition and sentence perception. This detailed information in the speech signal becomes part of the memory representation for spoken language. For instance, word identification performance has shown to be better for words that are produced by a single talker than for words produced by multiple talkers (Mullennix, Pisoni, & Martin, 1989), and sentence identification performance improved when listening to familiar as opposed to unfamiliar voices (Nygaard, Sommers & Pisoni, 1995).

Further experiments (e.g. Martin, Mullennix, Pisoni, & Summers, 1989; Goldinger, Pisoni, & Logan, 1991) have shown that specific details of the talker's voice are also encoded in long-term memory, including detailed information about speaking rate. Such results suggest that the acoustic features used to perceive the talker's voice are encoded into memory along with the linguistic message and form part of the neural representation of speech. Increased stimulus variability in an experiment may actually help listeners to encode items in long-term memory (Goldinger et al, 1991). Rather than discarding the rich characteristics of speech in favour of a highly abstract symbolic code like a string of phonemes, the human perceptual and memory systems appear to encode and retain very fine details of the perceptual event. The question remains whether this information is part of phonology, or whether it is stored elsewhere and accessed separately.

Pisoni (1997: 30) suggests that examplar-based or episodic models of categorization (for a review, see Goldinger, 1997) provide new solutions to the problems of invariance, variability, and perceptual normalization. According to multiple-trace theory, every stimulus, such as a spoken word, leaves a unique trace in memory (Goldinger, 1997: 33). Goldinger (1997) reviewed the details of a number of recent perception and memory experiments, showing that episodic memory traces of words contain indexical information about the speaker as well as the content of the linguistic message, and might therefore constitute the mental lexicon (there are several proposals for the episodic lexicon, some based purely on examplar traces, others suggesting the development of both lexical codes and episodic traces during speech perception). The theory is not without its problems, though, the most obvious one being the excess load on memory that the storage of episodes requires. Moreover, while support for the theory is available mainly from speech perception research, more research is needed to find out how traces are accessed during speech production.

1.5.4.2 Implications for language acquisition

If the observations concerning the kind of detail stored in the mental lexicon are true, then they suggest very close interactions between the form and content of the linguistic message and the listener's linguistic knowledge. All children encounter tremendous variation (of the type described in Section 1.5.4.1) in the language spoken around them. Such aspects of variability indicate that different speakers use different articulatory strategies in producing the 'same' lexical item and suggest that there may not be a single stored representation for a lexical item even for the same speaker, thus suggesting that a simple underlying form may be a cover for larger differentiated a set of traces. Pisoni (1997: 12) suggests that these types of variability are not only present in the acoustic signal generated in the utterance, but are also embedded in the articulatory and neuromuscular activity that generates the acoustic signal, and are therefore part of the speaker's competence rather than simply being a product of environmental factors. Thus, what a listener learns about a talker's voice, e.g. the acoustic correlates of gender, dialect, speaking rate etc. might be encoded and subsequently used to facilitate a phonetic interpretation of the linguistic content of the message (Pisoni, 1997: 10).

Since a multiple trace model (Hintzman, 1986; Jusczyk, 1997) allows traces of different phonetic tokens of a single item to be stored, children's increasing exposure to language will increase the range of variation that they recognise and accept. The ability to process and understand a novel or unfamiliar accent of one's own language is an important part of the individual's linguistic and sociolinguistic competence. The ability to accommodate a new accent into one's own speech patterns may also be important for the individual moving into a different dialect community (discussed in Section 1.5.3). The development of these abilities can be related to the developmental mechanisms responsible for the tuning of speech perception skills and for the gradual ability to understand regional accents (Nathan, Wells & Donlan, 1998: 363).

If the claims made by multiple trace models hold for speech perception and production, the implications for language acquisition suggest that children do not only develop abilities to discriminate and identify sounds, but they also learn to control the motor mechanisms used in articulation to generate precisely the same phonetic contrasts in speech production to which they have become accustomed in perception. By preserving very fine phonetic details and specific characteristics of the talker's voice, the developing perceptual system allows young children to accurately imitate and reproduce speech patterns heard in their surrounding environment. This provides them with a huge benefit in acquiring the phonology of the local dialect from speakers they are exposed to early in life (Pisoni, 1997: 28; Pisoni & Lively, 1995: 439), which will be discussed in more detail in Section 1.5.5.

1.5.5 Sociolinguistic competence

1.5.5.1. Introduction

Child language research in the 1960s was greatly influenced by Chomsky's views on linguistic competence and the dominance of Jakobson (1968), which highlighted universals in acquisition (MacNeilage, 1980). As a result, little attention was paid to variation of any kind within acquisition, particularly at the phonetic level. However, in the late 1960s, a growing number of researchers in sociolinguistics and psycholinguistics became dissatisfied with an idealised notion of competence. Linguists such as Labov (1966) began to pay greater attention to intra-language variation, and their proposals required broadening of Chomsky's view of what needed to be accounted for in language acquisition. Weinreich, Labov, & Herzog (1968) argued that it is unrealistic to view language as a 'homogeneous object'. They stated that 'native-like command of heterogeneous structures is not a matter of multi-dialectalism or 'mere' performance, but is part of unilingual linguistic competence' (1968: 101).

Over the years, many studies have documented the inherent variability in language, both in instances of language change and in cases of stable variation (e.g. Labov, 1963; 1966; Trudgill, 1974, 1986). Still, as Roberts (1997: 352) suggested, most studies have concentrated on the language systems of adult speakers and have rarely included speakers under the age of nine, so little is know about the acquisition of variation.

Since children acquiring language must obviously learn more than grammatical rules and vocabulary alone, other aspects of communicative competence deserve attention. An essential part of the communicative competence that children must acquire involves learning 'when to speak, when not, and... what to talk about with whom, when, where, in what manner' (Hymes, 1974: 277). The quote below is one of the earliest definitions of sociolinguistic competence, which has been documented as being part of the human endowment for several years.

> "Within the social matrix in which [a child] acquires a system of grammar, a child acquires also a system of its use, regarding persons, places, purposes, other modes of

> communication, etc.- all the components of communicative events, together with attitudes and beliefs regarding them. There also develop patterns of the sequential use of language in conversation, address, standard routines, and the like. In such acquisition resides the child's sociolinguistic competence (or, more broadly, communicative competence), its ability to participate in its society as not only a speaking, but also a communicating member." (Hymes, 1974: 75)

In acquiring full communicative competence, children must therefore learn to speak not only grammatically, but also appropriately. The socialisation process begins at birth. Children must learn the meaning ascribed to actions and feelings by their culture. Children participate in a variety of speech situations, with people who differ in age, sex, status, and familiarity, and whose speech will therefore vary in a number of systematic ways. Similarly, children learn the meaning of speech events and the socially acceptable or unacceptable contexts for those events (Lyon, 1996: 30). Therefore, learning to use language and learning to use language in context are inseparable. Some researchers suggest that children acquire sociolinguistic rules before they acquire structural language rules (Dopke, 1992). Andersen (1990) found that, when children aged 4 to 7 years are asked to take on different social roles (e.g. 'talk like a doctor/teacher/mother'), they vary their speech along a number of dimensions (e.g. register, words choice, syntactic devices).

In work on phonological acquisition, however, the majority of studies which have investigated phonological development have begun by asking how the child acquires the full inventory of adult oppositions (e.g. Ferguson & Farwell, 1975). Such a question originates from a phonemic assumption and the assumption that there is a stable model for the child to acquire. According to Local (1983), both these assumptions rest on shaky foundations, because they cannot enable us to investigate how children acquire the patterns of sociolinguistic variability reported for adult speakers (discussed below).

1.5.5.2 Sociolinguistic variability and phonological acquisition

Insights from recent variationist sociolinguistic studies of monolingual acquisition have shown that it is difficult to pinpoint a unitary system in any adult language and that different types of variability in the speech input that a child is exposed to such as dialectal, individual, and stylistic differences constitute part of the knowledge acquired by children. Studies by Foulkes, Docherty and Watt (1999), Docherty & Foulkes (2000), Local (1983), Roberts (1997), Roberts & Labov (1995), and Williams & Kerswill (1999) have embarked on methodologies inspired by the variationist work of Labov (1994; 2001) and the resulting advances in sociolinguistic theories (Chambers, 2002a). They highlight an aspect of phonological development that is normally overlooked in the majority of studies of children's speech, that of variability. These studies show that there often is no stable target model for the child to acquire, and that children acquire the range of sociophonetic variation that is acceptable in their speech community and the systematic distribution of the conditioned variants from a very early age as part of the development of their sociolinguistic competence.

Roberts (1997: 354) states that the rule-governed variation which has often been found to be part of the language of adult speakers is also part of the overall linguistic competence which a child must acquire in order to be a speaker of his/her language. Furthermore, as Andersen (1990: 32) notes, 'children must learn the dialect or set of dialects that will mark certain aspects of their social identity, including their region of origin, as well as their social class, ethnic group age, and gender'. In addition, they will learn the stylistic variation that will allow them to move from social group to social group, setting to setting, and conversational topic to conversational topic. Seeking out a model of language acquisition which denies the presence and the importance of the acquisition of heterogeneity is therefore considered unrealistic. Instead, a complete acquisitional model requires the inclusion of all forms of language, those which are variable as well as those which are categorical in nature. The quote below summarises Local's (1983: 452) view about what children do during the process of phonological acquisition.

> "It is clear that in the acquisition of phonology children must at least (i) sort out which parts of the variable input are linguistically relevant (e.g. the closed/open syllable patterns), which sociolinguistically motivated and which simply 'noise'. And having done this they must (ii) discover what range of variation under what conditions can be produced by them to count as 'hit'."

Local (1983: 449) considers the existence of a great amount of variation in children's phonology as a fact to be accounted for and not something troublesome to be cast aside when analysing data. He presents variation data from the speech of a Tyneside child and focuses on apparently 'trivial' details of phonetic variation which actually reveal important developments in the acquisition process of the vowel system of English. The analysis concentrates on the realisation of the stressed vowel in words such as *feet*, *cream*, *she*, *three*, in a Tyneside boy at the ages of 4;5, 5;0, and 5;6. The rule in question is the phonologically and morphologically conditioned vowel alternation in the lexical set corresponding to RP /iː/, which gives rise to a monophthong, typically [i], in closed syllables, as in *feet*, and a diphthong, typically [ɪi] in open syllables, as in *three*. Morphologically complex words such as *frees* have the diphthong, and therefore *freeze* and *frees* form a minimal pair.

The recordings consisted of sessions during which the child is interacting with a variety of interlocutors: his younger sister, peers, parents, other adults. There was considerable variability in the realisations of this vowel at 4;5, which were more variable than is to be found in the speech of the adults. However, some variants occurred only in particular phonetic/phonological environment. For instance, the [ɨ] variant only preceded polysyllabic words, while others occurred for particular stylistic purposes, such as [y], which is used for affective purposes and is part of what Local (1983: 451) calls 'crazy tonics'. These consist of variants that are restricted to talk during play and to talk which the boy's mother labels 'whingeing'.

The range of phonetic variation in the realisation of stressed /i:/ decreased as the child got older. More importantly, features that are particular to the Tyneside variety in given contexts increasingly became restricted to these contexts, e.g. closing diphthongs in open syllables and long monophthongs in closed syllables. Local (1983: 452) concluded that the child was still sorting out (i) the possible range of phonetic realisation for this stressed vowel, and (ii) the distribution in terms of syllable type, of the monophthongal and diphthongal variants. At the same time, the child was engaged in gaining control over the relevant localised phonological patterning of the variants of this vowel.

1.5.5.3 When does the acquisition of variation begin?

While Reid (1978) suggests that the acquisition of sociolinguistic variation takes place around the preadolescent stage, Romaine (1984) lowers the age to children as young as six years old, while Labov (1972) proposes an 'active period' for the acquisition of regional vernacular patterns between ages four and nine. More recently, however, developmental studies of several children are beginning to refine the schedule for childhood acquisition of adult norms (Foulkes, Docherty and Watt, 1999; Roberts, 1995; 1997).

Work by Roberts (1997) suggests that the acquisition of variation may begin during the preschool years when children are also acquiring the vast majority of categorical rules. Roberts investigated the acquisition of variation in Philadelphian preschool children (aged 3;2 to 4;11) by examining (-t, -d) deletion in consonant clusters in word-final position, one of the most well-documented variables rules in English. (-t, -d) deletion operates following internal and external constraints. Internal constraints include (i) grammatical constraints, such as the grammatical status of word containing the (-t, -d) segment, with the most favourable form for deletion being monomorphemic words in which (-t, -d) is part of the stem and contains no meaning apart from the words as a whole (e.g. *mist, nest*) and the least favourable form being the weak past tense (e.g. *missed*); (ii) phonological constraints, such as the identity of the segment following the (-t, -d) cluster, with obstruents being to most favourite and vowels and pauses being the least favourite.

External constraints that operate on (-t, -d) deletion include social class, gender, ethnicity, and conversational style (Roberts, 1997: 356). For instance, the phonological constraint related to the following segment was found to vary depending on the dialect in question, with Philadelphians favouring a following vowel for deletion more than a following pause, and New Yorkers favouring a following pause. Roberts (1997: 369) found that the children did indeed learn appropriate variation at an early age. They not only showed signs of having acquired the internal constraints on the (-t, -d) deletion rule, but also replicated the Philadelphian dialect pattern with respect to the following pause, which indicated that they were learning rules grounded in a socially transmitted dialect rather than being constrained by any universal principles or developmental factors.

Other research by Roberts & Labov (1995) has shown that children show signs of learning a variable rule that is still undergoing change in their community. Roberts & Labov (1995) investigated the acquisition of Philadelphian distribution of short /a/, which unlike (-t, -d) deletion, has been, and continues to be, involved in ongoing change, and is not consistent across a geographic area. The Philadelphian pattern of raising and fronting of short /a/ is very complicated (see discussion by Roberts & Labov, 1995: 102). Its complexity would seem to make it difficult to acquire, and Payne (1980) had found that for children aged eight to ten, who were acquiring short /a/ patterns as well as other Philadelphia variables as a second dialect (the speakers had moved to Philadelphia after having learned to speak another dialect). Payne concluded that, even if children were born and raised in the Philadelphia area, their chances of acquiring two systems were extremely slight unless their parents were also born and raised there.

Roberts & Labov (1995), on the other hand, focused on children aged 3 and 4, most of whom are in this ideal dialect learning environment, i.e. most parents were born and raised in Philadelphia. The children were found to have acquired short /a/ to a large extent, including, for instance, its tense realisation when preceding a nasal and a syllable boundary e.g. *sandals, mad,* and lax form in initial position if followed by a nasal and a vowel environment, as in *animal.* Moreover, the children showed signs of participating in the undergoing change that affecting short /a/ preceding /l/ and intervocalic nasals and actually pushing it forward (they produced more tensing in these two contexts than the adults, especially that the second context is not consistently prone to tensing by the adults).

Roberts & Labov (1995: 110) further suggested that the 3- to 4-year age level is a critical period for the acquisition of dialectal norms of the speech community, just at it is for language learning in general and, as recent research shows, for variation in particular.

Between the ages of 3 and 4, the children showed an increased adoption of the community norms. For instance, when short /a/ preceded /l/, the 4-year-olds were significantly more likely than the 3-year-olds to tense it, e.g. in *Sally* and *alligator*. The 3-year-olds did show tensing before /l/, but it was the 4-year-olds who were tensing short /a/ in this environment far more than the adults did. The children were actively learning the norms for the short /a/ system from the speech community, while at the same time participating in the lexical diffusion in progress. However, one 4-year-old child from the study, whose mother was not originally from Philadelphia, did not tense short /a/ as frequently as the other children in the relevant environments, which underlines the importance of the parents' dialect in providing an example for the children (discussed later). Roberts & Labov (1995: 111) concluded that the preschool period is a critical one for language learning, as it includes the acquisition of both grammatical rules and variable rules. Her findings emphasise the active participation of very young children in their speech community and the necessity of their inclusion in its complete description.

A recent study by Foulkes, Docherty and Watt (1999) shows that the acquisition of sociolinguistic variation might actually take place earlier than the above authors have suggested. The authors found signs of acquisition of detailed accentual features with regards to /t/ production by young Tyneside children aged 2-4 (/t/ realisation in the adult community is very complex and exhibits a strong correlation between a number of variants and social factors such as age, sex, and social class). Moreover, Foulkes, Docherty and Watt (1999) showed another strong correlation between these children's productions and an innovative form of pre-pausal /t/ in Tyneside English, mainly a pre-aspirated variant (e.g. [naɪʰt] 'night'), which is associated with young working class women (same demographic group as the children's mothers). Such findings confirms Roberts & Labov's (1995) claim that very young children actively participate in their speech community, including learning features undergoing changes. The authors concluded that the children were not discarding features that do not have an essential linguistic function, i.e. that are not essential for lexical contrast. Language- and accent-specific features that are not necessarily distinctive also form part of what a child acquires about the sound patterns of their native language and the child may not originally distinguish between the two. Moreover, given the great amount of variability that characterizes the input that children receive, it is not surprising that they might reach different conclusions about the adult patterns and in turn contribute to language change in their society (Kerswill, 1996).

1.5.6 Role of the environment in the acquisition of sociolinguistic competence

Children's social identities develop rapidly from infancy to post-adolescence, as they pass through the 'life course' (summarised in Giddens, 1989: 82-85), from a strong attachment to the caregiver, through the transitional stage of adolescence in which individuals see a conflict between their wish to behave as adults and their treatment by the adult world as children, to the boundary of adulthood with the need to make independent life-decisions. Each stage is mirrored by difference in language use associated particularly with the child's orientation to other people. Starting from a parent-centred orientation, the young child expands his/her range of social contacts to other, often older children, leading to distinctive teenage peer groups with their attachment the youth culture and opposition to adult norms (Eckert, 2000; Kerswill, 1996). The growing child's changing orientation to different groups of people must be seen as part of his/her maturing sociolinguistic competence. Two such stages will be discussed in this section: the period from birth till the child is six years old, as it is marked by influence from the parents, and the period from age six to twelve, as it marks the transition to influence from peers.

1.5.6.1 Influence of caregivers on infants and young children (age 0 to 6)

According to Kerswill (1996: 190), children acquire most, if not all, the phonological features of their local variety by the age of 6. There seem to be three types of features. First, there are those that must be acquired much earlier than 6, if they are to be acquired at all, including those with lexically irregular conditioning. Second, there are those features with morphological conditioning which may be acquired up to the age of 6 through prolonged exposure to speech community members (especially caregivers). Finally, there are phonologically simple features which can be acquired at any time. Whether any of these features are acquired or not is also dependent on the availability of other models and the social evaluation of the features.

In this first stage of language acquisition, most input initially comes from the main caregiver, with the role of other adults and children gradually increasing (Kerswill, 1996: 181). It is at this stage that we might find the initial transmission of dialect features and sociolinguistic competence. A lot of researchers have noted the importance of parental input for the acquisition of dialect features by the child at this stage, especially with regards to complex features or features that are undergoing change. For instance, Payne (1980) found that out-of-state children moving into Philadelphia did not acquire the 'correct' patterning for the tensing and raising of the short /a/ vowel, which is both phonologically and lexically determined, although they had acquired low-level phonetic rules and realisations of other vowels, even for those children who had arrived by the age of eight (Payne, 1980: 124). The short /a/ patterns were only acquired by those children whose parents had the Philadelphian patterns themselves. Similarly, Trudgill (1986: 35) claims that the acquisition of /u:/-/ʌu/ distinction in Norwich requires that at least the speaker's mother must have a local accent.

Chambers (1973) examined the acquisition by Canadian children of the Southern British English opposition between /ɔ:/ and /ɒ/, absent in their own dialect. Of these children, only those who arrived in England at a relatively early age (in this case before the age of 13), made any progress in separating the two lexical sets. This age is late compared with the age that appears to be critical for the Philadelphia short /a/, although this probably reflects differences in the difficulty of the particular features. Chambers' data allowed him to propose a 'critical age of dialect acquisition' and place it somewhere between the ages of 7 and 14, beyond which complex rules and oppositions are rarely acquired as second dialect features.

'Simple' rules, described by Kerswill (1996: 187) as rules that have no phonological, morphological, or lexical constraints, might be acquired with or without influence from the adult model. Kerswill gives the example of the variable (ou), which refers to the fronting of the offset of the diphthong /əʊ/ in southern British English. Kerswill & Williams (2000) studied children's and adults' speech in the New Town of Milton Keynes. 48 children (from each of three age groups: 4, 8, and 12) who had been born in Milton Keynes or who had moved there within the first two years of their life were recorded in various tasks and their parents were interviewed. The (ou) variable was investigated in the children's and adults' speech, and has shown a good deal of fronting by all older children aged eight and twelve, but not all the four-year-olds. For this youngest group, three strategies were detected. First, there might be accommodation to the majority of (older) Milton Keynes children. Second, children can model themselves on one or the other parent. Third, children can strike a compromise between their parents. By age six, one child had abandoned the Scots vowel system that he had acquired from his father, which shows that a feature like /əʊ/ fronting can be acquired as a second dialect feature by older children.

There are also cases where there is no clear link between the adult and the child feature. For instance, Kerswill (1994) compared the use of [f] for /θ/ and the labiodental [ʋ] for /ɹ/ in six 4-year-olds and their principal caregivers. For the (θ) variable, all six children used [f], while only two of the caregivers did. From a sample of 48 children (aged 4, 8, and 12) the children's percentage use of [f] was 68.3, with all the children using at least some [f], while the percentage for the caregivers was 17.0 (Kerswill, 1994: 18). For the use of [ʋ], the picture was more complex in that not only articulations were more variable, but they also varied along on a phonetic continuum. Since these changes, along with many non-developmental changes, are spreading throughout other regions in the UK, Kerswill (1996: 189) maintains that the persistence of immature forms cannot be considered as the cause of change; it may simply be a facilitator.

1.5.6.2 The influence of peer groups on preadolescents (age 6 to 12)

At the preadolescent age, most areas of language are assumed to be fully mature, with the exception of the command of an adult range of speech styles (Kerwill, 1996: 191). Chambers (2002a: 176) notes that certain complex rules and new oppositions can still be acquired by this age group, although the children's abilities are beginning to be restricted as they reach their 'critical age of dialect acquisition'. Similarly to the findings in Section 1.5.6.1, there are limits to vernacular acquisition at this stage as well, unless caregivers themselves speak something close to it (e.g. Hewlett, Matthews & Scobbie, 1999).

As mentioned in Section 1.5.6.1, phonologies are more or less acquired by this age, but children begin to assert themselves outside the home. Although the family normally provides the first speech model for infants, within a few years it is replaced by a more significant one, that of friends (Chambers, 2002a: 175). Schoolchildren normally speak more like their peers than like their elders, as classmates and close friends are linguistically (as well as in other ways) more influential than teachers and parents. For children whose parents belong to a different speech community from the one in which the children are being raised, the predominance of peers over elders becomes most obvious. The children have two models of dialect transmission, but one of these models, the parents, is never chosen in normal circumstances. This is said to take place regardless of whether the parents speak a different dialect, or a different language (Chambers, 2002b: 175).

Chambers (2002b: 175) gives the example of monolingual Italian adults immigrating to Sydney who learn English and end up speaking an ESL dialect, although their Australian-born children do not acquire their parents' ESL dialect (more on this in Section 1.5.5). Similarly, when Scots school teachers settle and live in London, they retain most features of their native Scottish accent, although their London-born children do not retain their parents' Scottishness in their own accent. The children might be expected to have some Scots dialect features, at least up to the age of around five, and should lose them, probably rapidly afterwards. There are therefore subtle patterns of dialect shifting, comparable to style-shifting in unmixed dialect situations, as the children make the transition from parental norms to their peers (Chambers, 2002b: 176).

In looking at (ou) fronting again, a feature of younger speakers in the southeast, Kerswill & Williams (1994: 20) found a difference across the children, with the younger ones producing patterns that were closer to their caregivers, and the older ones moving away to a new norm, for which the model is partly external (/əʊ/ fronting is found throughout the southeast). No correlation was found between caregivers' and children's fronting except for two 4-year-olds with non-southeastern vowels. For the 8 - and 12-year-olds, the mothers' pronunciation of this vowel had no effect on their children, and the main factor was the child's orientation towards the peer group. High scorers were described as out going, with friends in the school, and therefore very well integrated into a group of friends, while low scorers were described as shy, loners, and having few contacts with family members apart from their parents. It is the sociable and peer-oriented children who are in the lead in this change (Kerswill & Williams, 2000: 236). This suggests that the children were focusing on a norm that is different from that of the adults, and, in this New Town context, this constituted evidence for a new variety.

Hewlett et al (1999) examined the acquisition of the Scottish Vowel Length Rule (SVLR) and 'Voicing Effect' (VE) in seven Scottish English speaking children aged six to nine years and who were born and bred in Edinburgh. VE is found in most English accents and concerns vowel duration; vowels tend to be longer before a voiced consonant than before a voiceless consonant e.g. 'bead' (long vowel) versus 'beat' (short vowel). SVLR is found in Scottish English and states that some vowels are long in open syllables and before voiced fricatives, /r/, and a morpheme boundary, and short elsewhere. There is therefore an overlap between the two rules, as the length distinction between 'bruise' (long vowel) versus 'Bruce' (short vowel) results from the application of either rule, while 'brood' is predicted to differ (short according SVLR, and long according to VE). The study aimed at examining to what extent VE and SVLR operate in Scottish children's speech, and whether there are any differences according to parental accent. Two of the children had two Scottish parents, two had one Scottish and one English parent, and the remaining three had parents neither of whom spoke Scottish English (but spoken another variety of English).

There was a difference between the children from households in which either or both parents spoke Scottish English and those children from households in which neither parent was a Scottish English speaker. Children from households in which either both parents or one parent spoke Scottish English showed a strong SVLR effect and a very modest VE effect (Hewlett et al, 1999), as they produced lengthening before a fricative that was at least 52% more than lengthening before a stop. On the other hand, the three subjects neither of whose parents were Scottish English speakers showed a more pronounced VE and a modest SVLR effect.

Although the three children were judged to speak a Scottish English accent, their detailed implementation was 'less Scottish' in showing an influence from accents containing a strong VE and no SVLR. Two of these subjects actually had a higher percentage of vowel lengthening in VE contexts than in SVLR ones. The third subject had an intermediate pattern, with a VE that is modest in comparison with the other two subjects but strong in comparison with the children of Scottish parents. The authors noted that similar cases of intermediate values between two alternative phonetic targets are attested in the area of L2 learning (see Flege, 1995) and suggested that something similar may happen in the case of first language acquisition where there is significant exposure to two accents. 'Phonetic implementation does not necessarily require an all or none choice between two accents and the phonetic system may fall midway on a continuum between the two' (Hewlett et al, 1999).

1.5.6.3 Summary

In sum, studies reviewed in Sections 1.5.5 and 1.5.6 suggest that children acquire their first language by detecting the patterns underlying variability in phonology, morphology, and syntax. This variability may simultaneously encode both sociolinguistic and linguistic information, and it is several years before the child has successfully analysed the complex patterning of the adult model. Exactly when a child acquires a feature of his or her first dialect depends on the linguistic level, the complexity of the conditioning environment, and the child's age (Kerswill, 1996: 198). All these issues have several important implications for bilingual acquisition, which we now turn to.

1.5.7 Implications for bilingual language development

1.5.7.1 Early perception/production abilities

With respect to early perceptual abilities, studies suggest that infants who are exposed to two languages from birth tune into the fine-grained phonetic structure of their languages from a very early age and use their acquired perceptual knowledge for later production. Since recent research has also shown that infant perceptual abilities are not suddenly lost after the first year, this suggests that successive bilinguals can also acquire the phonetic patterns of their second language through extensive exposure.

As for production abilities, early productions of monolingual children show a great amount of variation within and across individuals. This variation is partly due to articulatory maturation and the ongoing change in the child's production and perception, but also due to the varied input that the children receive. With respect to articulatory maturation, it is important to learn about monolingual patterns of development before interpreting bilingual behaviour. What might be considered as an influence between the two languages in the bilingual's production or a delay in bilingual language development might actually be part of normal monolingual developmental patterns. Monolingual developmental patterns are in turn extremely varied, which underlines the importance of considering individual differences before judging a bilingual's production as 'deviant' or compliant with the 'norm'. Watson (1995) points out that 'it is difficult to be sure that a bilingual is doing something that a monolingual would never do', as monolingual norms are themselves constructs and they may conceal enormous variation.

1.5.7.2 Input variability

Issues in monolingual development that were reviewed in this chapter point to the difficulty in pinpointing a unitary system in any adult language, as adult input varies according to speakers, dialects, speaking rate, speaking style, etc. This varied input has been shown to constitute part of the knowledge acquired by children. On the one hand, variability in the input leads to idiosyncratic ways of interpreting the adult forms by the child and to noticeable individual differences in phonological development. On the other hand, exposure to structured variability enables the child to acquire the range of sociophonetic variation that is acceptable in their speech community.

The bilingual child presumably receives more varied input within and across the languages that are spoken in his/her environment. This input often varies between standard, non-standard, and non-native varieties, depending on the speakers and contexts that the bilingual is exposed to the most. In research on bilingual development however, the issue of variability in the input has rarely been considered. Studies on bilingualism rather treat the two systems involved as homogenous entities that the bilingual child should be aiming for. It is vital to take input variability into consideration before analyzing bilingual speech. First, what might be considered an influence from one language onto the other or a sign of the child initially using one system might actually be due to characteristics of the input that is available to the child. Second, if we can find evidence that a bilingual child has acquired fine grained details of sociolinguistic relevance to the community in which it is growing up, this would offer very useful information about the phonological knowledge (in the broadest sense) it has acquired. Studies on bilingualism have on the whole ignored such aspects in acquisition and instead concentrated on evidence for the successful acquisition of contrasts by the child. Evidence may also come, however, from the investigation of detailed language- and accent-specific phonetic/phonological features, since these are important to the bilingual in developing a sociolinguistic identity. Variability shapes the bilingual's acquisition process and influences his/her choices during the course of accent acquisition. Therefore, when approaching phonological aspects of acquisition, it is important that a thorough assessment be made of variable targets a child must aim for in order to speak like a mature member of its community.

1.5.7.3 Role of input in bilingual situations

While for monolingual children one or both parents often provides the initial model for the phonological features of the local variety, the bilingual child may or may not have this initial model available depending on whether either of the parents speaks the language/dialect of the community.

Many existing case studies on bilingual acquisition describe a situation in which the parents have different native languages, and each parent addresses the child in his or her language (e.g. Leopold, 1947; Ronjat, 1913; Schnitzer & Krasinski, 1994;1996; Volterra & Taschner, 1978). This has come to be known as the 'one-person-one-language' approach. It is sometimes claimed to be the 'best' way to bring up a bilingual child successfully, though the evidence is questionable and there are many different situations in which children acquire two languages. One reason that is given for recommending the 'one-person-one-language' to parents is that it will allegedly help the child to distinguish between the two languages in her environment if she hears them from two separate people. However, as Romaine (1995) and Genesee (1989) suggest, the mixed-language input is more common than it seems in the literature (each parent may use more than one language with the child, sometimes across utterances and other times within the same utterance). Moreover, Deuchar & Quay (2000) have shown that language differentiation by the child does not depend on associating one language with one person. In their study, the child was addressed by both parents choosing their language according to their location (at home or not) and according to the presence (or not) of monolingual speakers of English.

Apart from the one-person-one-language situation, there are several types of bilingual families, ranging from ones with both parents being bilingual to ones where both are monolingual but where the children later become bilingual (Hoffmann, 1991: 46). As a result, the concept of parental input varies with respect to the language(s) the parents speak, their attitudes towards either, and the strategies they adopt to enhance or inhibit the chances that their children will become bilingual. The type of family will no doubt affect the degree of bilingualism of the children.

Parental attitudes towards the language(s) they/their children speak usually depend on (i) the family's degree of integration into society; (ii) the status of the languages that the parents wish to pass on to their children (e.g. minority versus majority language); and (iii) the reasons that would motivate them to encourage the acquisition of the language(s) by their children (Lyon, 1996: 35). In bilingual communities, a considerable proportion of the families speak two languages, and bilingualism may be considered as the norm. On the other hand, there are communities where immigrant families become bilingual in the host country but remain a minority in terms of the parents' first language (as in the case of the subjects in my study). In this case, the maintenance of the children's first language is usually the responsibility of the family, while the medium of instruction in school is the language of the host country (Lyon, 1996: 45). The way in which these children become bilingual is unpredictable and depends on family circumstances (visits abroad, purchase of books and videos in the minority language, bi-literacy, etc.) and the surrounding social situation affecting them (whether there are other migrants of the same origin).

The world of nurseries and later schools, playground and neighbourhood introduces the bilingual child to considerable social influences that are quite different from the ones experienced at home. In the case of bilingual second language acquirers, the second language often becomes the principal language of the child even in cases where both parents speak the first language; the community language is needed for social contacts and for being able to follow the school curriculum (Hoffmann, 1991: 46). From this stage onwards, the socialisation process becomes very complex, as the bilingual child trying to establish an identity has two models in his/her community instead of one (Hoffmann, 1991: 148). The degree of success of growing up bilingual depends on (i) whether the child has enough exposure to the language used only in the home; ii) whether (s)he considers it important to use both languages; and (iii) whether there is social support for the minority language (local centres where the minority language is used or taught). Such resources are usually less available in cases of individual bilingualism as opposed to societal bilingualism, but motivation and parental support play an important role as well.

Kravin (1992), describing her Finnish-English bilingual child's experience, addresses the issue of whether an L2 parent can provide enough input in isolated settings, when (s)he is the only source of input and the child's only feedback (she was the only source of Finnish to her son while the father was a monolingual English speaker, as the family lived in the US). Even though a strict one person-one language approach was followed with the author's son, Dan, he did not acquire both languages, and parental input did not determine the course of language development. One factor that does need to be taken into consideration, though, is that the mother spoke English and not Finnish to the father, and this may have had an influence on Dan's development too. As Dan grew older, he focused on the surrounding social and linguistic environment, and mother-child relationship slowly diminished in favour of the peer relationship. Therefore, although maternal input is so close a relationship, it cannot by itself insure the acquisition of a language. Domains that the child mastered in English (conversing with peers and adults, watching TV, etc.) became difficult for him to express in Finnish, which made him more cautious in using the language and more frustrated about not being able to communicate about his favourite topics with Finnish peers. Kravin (1992: 323) concludes that every child is able to become bilingual on a native level, but if the input provided for this development is restricted to only parental speech, development stagnates and later attrition might occur.

Little is known about what happens in bilingual situations upon the acquisition of a given dialect by the child, especially if the second language is acquired outside the home. In such cases, the parents are usually immigrants who end up speaking a second language dialect that includes features of interlanguage, but their children will not speak the parents' dialect. As in monolingual situations, people of the same age will presumably have more influence on how the children sound and how they use the language than their parents do (Hoffmann, 1991: 26). However, what will be missing is the initial parental model that should provide the child with the basic phonological and sociolinguistic patterns of the dialect. Hardly any investigation has considered the particular details of such a situation and the resulting sociolinguistic choices available for the bilingual child. In cases where the home language of the child is different from that of the society, sociolinguistic competence in both languages is difficult to achieve, since socio-cultural upbringing in both cultures, a pre-requisite of sociolinguistic competence, is prone to be uneven. It requires considerable effort on the part of the parents as well as strong links with the speech community of each of the languages (Hoffmann, 1991: 26).

Chambers (2002b: 121) suggests that the children of immigrants will have an innate 'accent filter' that is part of their sociolinguistic competence. The existence of the accent filter follows from what Chambers calls the 'Ethan Experience', named after the son of eastern European immigrants in Toronto. Ethan's parents are fluent and articulate ESL speakers with (by their own admission) 'medium-to-strong accents'. Ethan was born and raised in Toronto, and spoke English with the same accent as all his native-born classmates and not at all like his parents. Even as a pre-schooler, Ethan never acquired his parents' accent-features, not even in isolated words. Chambers claims that this fact is not unique to Ethan, or unusual, but is rather so common that it usually goes unremarked.

Ethan, at no time, even momentarily, acquired pronunciations with tapped /r/ or close versions of lax vowels, characteristic of both his parents' English pronunciations. Chambers (2002b: 121) claims that this holds equally for countless other children, which suggests that it is principled behaviour that needs to be accounted for in a theory of language convergence. Chambers goes on to describe how Ethan comes equipped with an innate filter so that when he hears his mother say 'cherry' with tapped /r/, he hears it as retroflex and pronounces it that way. When he hears his father say a word like 'cell' with the tonic vowel pronounced [e:], he hears the vowel as [ɛ], and says it like that. These filterings apparently take place beneath consciousness. Ethan was well into his school years when he realised that his parents' English was foreign-accented. This too, Chambers claims, is typical of children growing up in households where the parents are fluent ESL speakers. Due to the efficiency of the innate accent filter children presumably do not ignore the foreign-accent features in their parents' speech, but simply fail to hear them and end up acquiring the native accent of their peers (Chambers, 2002b: 122).

While Chambers is right in observing that many children of immigrant families end up speaking more like their peers than the second language accent of their parents, some of his assumptions are debatable. With respect to the claim that children fail to hear their parents' foreign accent and to recognise it as different from theirs, there are countless anecdotes about bilingual children showing awareness of their parents' L2 accent and sometime 'correcting their mistakes'. As for the assumption that the bilinguals will end up speaking like their peers, studies reviewed in Section 1.5.6 show that some phonological features will not be acquired by the child unless one of the parents speaks the dialect of the community. Therefore, despite the importance of peer influence, parental influence is also essential for the acquisition of complex features that requires extended exposure. Chambers (2002b: 123) sees the Ethan Experience as a tool which allows individuals to separate themselves from certain unconventional communities that fall within their worlds, presumably in order to allow them to participate fully in the communities that will play more integral roles in forming their identities. Chambers is therefore describing only one possible scenario for bilingual families, which involves the children's will to fully integrate within the host society. There are many other possible outcomes; therefore, understanding the social context of the bilingual acquisition is essential in many ways, as it leads to a better understanding of the language input to the child, the attitudes fostered by the family towards either language, and the resulting effect on the child's linguistic behaviour.

1.6 Other methodological issues

1.6.1 Later bilingual development

While the one-or-two system discussion concentrates on early stages of acquisition, little is known about what happens in the language acquisition process later; most studies end when the child reaches his/her second or third birthday and differentiation is no longer an issue, as it is taken for granted.

Even if the findings suggest that the children's phonological systems are indeed differentiated, the question still remains whether these systems are entirely autonomous and like that of 'two monolinguals in one', or whether some crosslinguistic influences between the two systems are apparent (Paradis, 2001: 21). Research on adult bilinguals has shown that interactions between a bilingual's two languages is apparent on many levels, suggesting that this third possibility is a common outcome of bilingual development (e.g. De Groot, 1993; Hazan & Boulakia, 1994; Paradis, 1997). The interactional perspective on language representation in adult bilinguals is briefly summarised by Grosjean (1995: 259): 'Bilinguals are not the sum of two complete or incomplete monolinguals but have a unique and specific linguistic configuration'. Therefore, the 'one system or two?' dichotomy posed in much research on child bilinguals may be too simplistic; if adult bilinguals never achieve full separation of their systems on all levels, then it may be inappropriate to expect child bilinguals to do so. It may be more appropriate to approach the study of bilingual language development with the expectation that interactions between the two languages will occur, even after differentiation (Paradis, 2001: 21). However, other types of language interaction have often been interpreted as evidence for transfer or interference (deviation from the language being spoken due to the influence of the other language), without regard to the context or the language mode (discussed in Section 1.6.3) in which the so-called 'interference' took place.

Paradis & Genesee (1996: 3) define interdependent development as 'the systemic influence of the grammar of one language on the grammar of the other language during acquisition, causing differences in a bilingual's patterns or rates of development in comparison with a monolingual's. Interdependence could take several forms. First, the simultaneous acquisition of two languages might pose challenges to the faculty that could result in slower development for bilinguals in comparison to monolinguals acquiring the same languages. While this might influence general development in both languages, it could also influence development of specific aspects of one or both languages. Second, and in contrast, bilingual acquisition could accelerate language development in case where two language share certain structural properties and especially structures that normally emerge earlier in one of the languages when acquired monolingually. For example, monolingual French-learning children acquire finite verb forms at an earlier age than monolingual English children (Paradis & Genesee, 1996). For this reason, children acquiring English and French simultaneously might show an accelerated acquisition of finiteness in their English as a result of the early emergence of finiteness in their French verbs. Yet another interdependence is transfer, the systematic incorporation of a linguistic property from one language into the other (Genesee, 2001: 159). Transfer results in a deviant structure or pattern in comparison to the target language, for e.g. English-French bilinguals placing negatives after lexical verbs in English and French.

Note that transfer does not constitute evidence against the autonomous development of the linguistic systems of bilingual children. Transfer may be temporary. Given sufficient exposure to two languages, bilingual children can acquire the same grammatical competence in each of their languages as monolinguals in the long run (White & Genesee, 1996). Therefore, whether or not individual children exposed to two languages at birth become bilingual is largely a matter of circumstance rather than inherent limitations in the language faculty's ability to handle two languages at the same time (Genesee, 2001: 160). Moreover, monolingual children also exhibit non-target constructions that deviate from adult forms but nevertheless conform to the overall structure of the target language (e.g. 'goed' instead of 'went'). But while the non-target constructions of monolingual children are always based on the same system in which they occur, bilinguals can cross language boundaries. The language faculty is therefore able to coordinate different linguistic systems in the course of development (Genesee, 2001: 161).

1.6.2 Language interaction

As mentioned in the previous section, the competence of bilinguals cannot be considered as the sum of two linguistic codes, nor can it be measured in terms of monolingual standards; the linguistic experiences that either group encounters are not directly comparable and take place in different environments. This does not suggest that monolingual children are at an advantage in terms of their ability to master a linguistic code. There are speech strategies that are unique to bilinguals and which are used as aids to communicative ability. These are known as code-switching and code-mixing and are often reported in the literature describing bilingual performance, although they have been used with widely different meanings (De Houwer, 1998: 252).

1.6.2.1 Code-switching *versus* code-mixing

Until the 1960s, the literature had a tendency to label all language contact phenomena observed in bilingual production as interference (see Weinreich, 1953). More recently, due to extensive research on code-switching, traces of language contact have often been tied together under the general term 'code-switching'. Grosjean (1995: 263), however, suggests that at the level of the bilingual's underlying linguistic systems and the psycholinguistic processes that take place during the perception, comprehension, and production of language, we may be dealing with different phenomena which, on the surface, may appear at times to be identical.

However, attempts to distinguish between the two have resulted in conflicting positions about the characteristics of each label. Code-switching is often defined as the alternate use of two languages or linguistic varieties within the same utterance or during the same conversation (McLaughlin, 1984). Code-switches normally take place across phrase or sentence boundaries. Code-mixing, on the other hand, is defined as the insertion of a single element, or of a partial or entire phrase, from one language into an utterance in another (Hoffmann, 1991: 105). Elements can be phonological, morphological, syntactic, lexico-semantic, phrasal, or pragmatic. Other researchers distinguish between 'code-switching' and 'code-mixing' depending on the bilingual's perceived ability to select the language according to the interlocutor. Redlinger & Park (1980: 337), for example, use the term 'language mixing' to refer to the "bilingual's indiscriminate combination of elements from each language", suggesting that mixing takes place when the person is not able to differentiate between the two languages. 'Code-switching', on the other hand, reflects a bilingual person's ability to select languages according to the interlocutor or the situational context, and is therefore seen as a sign of pragmatic competence. Of course, not all researchers agree with this definition, especially those who have provided evidence for the early differentiation of the two languages by bilinguals. Research on both adult bilinguals (e.g. Myers-Scotton, 1997) and child bilinguals (e.g. Meisel, 1989; Paradis, Nicoladis & Genesee, 2000) indicates that their intra-utterance code-mixing is not random, but is grammatically constrained and, furthermore, complies with language-specific characteristics of the participating languages.

1.6.2.2 Code-switching *versus* borrowing

According to Grosjean (1995: 263), language mixing is made up of two different processes: code-switching and borrowing. 'Code-switching' is shifting completely to the other language for a word, a phrase, a sentence, etc. 'Borrowing' is taking a word or short expression from the other language and (usually phonologically or morphologically) adapting it to the base-language. However, the distinction between the two is not always straightforward, and code-switching and borrowing are still the focus of much current controversy. Some researchers argue that these language contact phenomena should be distinguished (e.g. Muysken, 1995; Poplack & Meechan, 1995), though consensus is yet to be reached on which surface manifestations should be classed in which category. Others contend that code-switching and borrowing are either undifferentiated in the bilingual speaker or effectively indistinguishable (e.g. Myers-Scotton, 1997).

Poplack & Meechan (1995: 201) note that the root of the problems stems from identifying the status of 'lone' words which are inserted from one language into the other and which, ironically enough, constitute the richest portion of any bilingual corpus studied. In many communities, single word switching is the commonest kind; these are often considered loans rather than code-switches. However, the distinction is not always reliable (Gardner-Chloros, 1995: 72). First, both loans and code-switches can be morphologically and phonologically integrated or un-integrated with the surrounding language, depending on a wide variety of personal and linguistic factors. Second, there are examples of both loans and codes-switches filling 'lexical gaps' in the surrounding language and of them adding themselves as a further option to the 'native' equivalent. Third, although loans are often nouns, all grammatical categories are potentially 'borrowable'; conversely, in many contexts, noun code-switching is the most common kind to occur in the data. One explanation for the lack of reliable distinction is that every loan starts off life as a code-switch, and some of these code-switches become generalised and spread through the community (Gardner-Chloros, 1995: 72). Like other linguistic changes, they spread irregularly, and in some sub-sections of the community they acquire status sooner than in others.

Bilinguals usually explain that the reason they code-switch is that they lack the facility in one language when talking about a particular topic (Grosjean, 1982: 150-157). They report that they switch when they cannot find an appropriate word or expression or when the language being used does not have the items or appropriate translations for the vocabulary needed. Some notions are just better in one language than the other. At other times, the bilingual simply has not learned or is not equally familiar with the terms in both languages. Bilinguals might know that with more effort and time, they could find the appropriate word or expression in the base language, but may claim to be tired, lazy, or angry, and therefore resort to the most available word. Code-switches may be also used for a particular topic (money issues, geometry, etc.), or may involve fixed phrases or greetings or parting and discourse markers such as,-'you know' and 'pero' (*however*). Code-switching may also be used to convey semantically significant information (e.g. reflecting personal involvement or detachment), or emphasising a point (e.g. terminating an interaction, underlining a request, etc.). Switching can also be used to signal group solidarity or exclude someone in a conversation. It can be used to quote what someone said, to raise one's status or give them added authority.

Studies show that adult code-switching is a sophisticated, rule-governed communicative device used by linguistically competent bilinguals to achieve a variety of communicative goals such as conveying emphasis, role-playing, or establishing a socio-cultural identity (Genesee, 1987: 164). It has highly structured syntactic and sociolinguistic constraints (so that the rules of both languages are respected). Adult bilinguals also switch between the two languages as a function of certain sociolinguistic factors, such as the setting, tone, and purpose of the communication or the ethnolinguistic identity of the interlocutor. According to Myers-Scotton & Jake (2001: 86) bilingual speakers take into account attitudes towards the linguistic varieties that the speakers have the potential to employ (i.e. sociolinguistic considerations); they also take into account perceptions of their own proficiency and the proficiency of the interlocutor in the same linguistic varieties (i.e. psycholinguistic considerations). Possible attitudes towards even specific varieties of bilingual production are also weighed.

1.6.2.3 Code-switching *versus* interference

Early researchers on bilingualism have proposed definitions of interference that are based on research on second language acquisition (e.g. Albert & Obler, 1978; Selinker, 1972). Transfer during second language acquisition takes place when the learner imposes structures on the new language which (s)he transfers from the previous language or languages. This basic acquisitional strategy is considered to result in erroneous productions in the second language in those instances in which the rules or structures of the second language do not coincide with the rules or structures of the first language (Albert & Obler, 1978: 209).

In a bilingual context, Weinreich (1953) defines interference as 'those instances of deviation from the norms of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language'. Haugen (1956) refers to it as 'the overlapping of two languages'. Mackey (1965) defines it as 'the use of features belonging to one language while speaking or writing another', and Clyne (1972) calls transference 'the adoption of any elements or features from the other language'.

However, as Grosjean (1982) notes, those definitions do not distinguish the controlled and more or less conscious use of code-switching and speech borrowing from the involuntary or accidental use of elements from the wrong language when speaking to a monolingual. Grosjean (1995: 262) distinguishes between 'static interferences' which reflect permanent traces of one language on the other (such as a permanent accent), and 'dynamic interferences', which are transient intrusions of the second language (as in the case of the accidental slip on the stress pattern of a word due to the stress rules of the other language). The confusion between the two has led certain researchers to reject the term 'interference' because it carries pejorative and disruptive connotations, and has led to stress the positive aspects of interference in a bilingual environment or the use of other terms such as 'interaction'. Haugen (1977: 322) writes:

> "we need to get away from the notion of 'interference' as somehow noxious and harmful to languages. The bilingual finds that in communicating he is aided by the overlap between languages and he gets his message across by whatever devices are available to him at the moment of speaking."

1.6.3 Language mode

The concept of 'language mode' was developed by Grosjean (1982; 1998; 2001) and refers to "a state of activation of the bilingual's languages and language processing mechanisms at a given point in time" (Grosjean, 2001: 3). Such mode operates along a continuum ranging from monolingual (i.e. one language is activated) to bilingual (i.e. both languages are activated) with intermediate modes in between depending on factors such as who the bilingual is speaking or listening to, the situation, the topic, the purpose of the interaction, and so on. At one end of the continuum, the bilinguals are at, or close to a monolingual mode in that they are interacting only with monolinguals of one or the other of the languages they know. One language is active and the other is only very slightly active. At the other end of the continuum, bilinguals find themselves in a bilingual language mode in that they are communicating with bilinguals who share their two (or more) languages and where language mixing may take place (i.e. code-switching and borrowing). In this case, both languages are active but the one that is used as the main language of processing (the base or matrix language) is more active than the other. These are end points, and bilinguals also find themselves at intermediary points depending on the factors mentioned above (Figure 1.1).

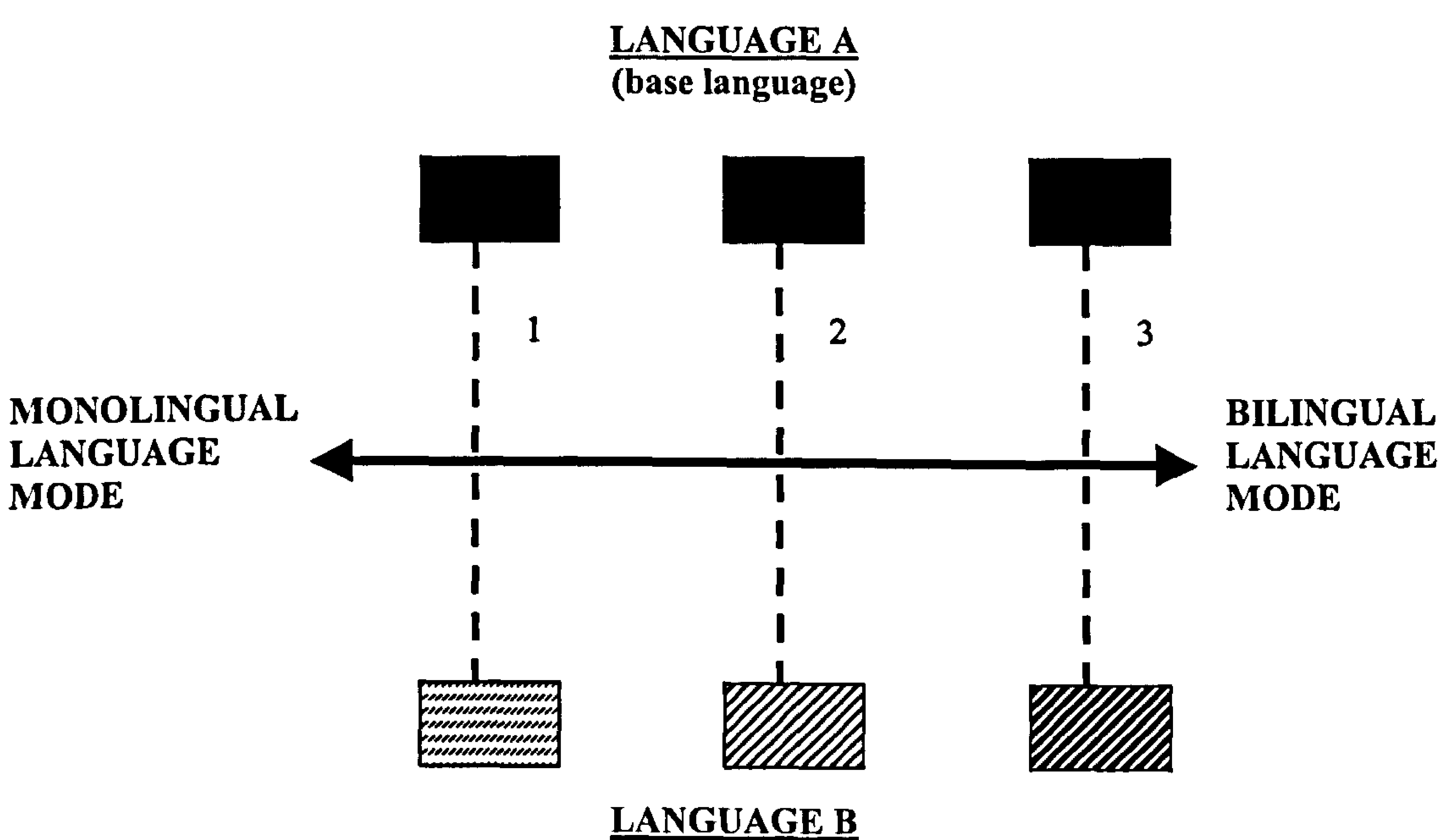


Figure 1.1: Visual representation of the language mode continuum as presented in Grosjean (2001: 3).

Three hypothetical positions for the same bilingual are presented in Figure 1.1. In all positions, the bilingual is using language A as the main language of communication (the base language) and it is therefore the most active (black square). In position one, the speaker is at, or close to a monolingual mode: language A is totally active whereas language B is deactivated (most often unconsciously) so that it does not lead to miscommunication. This mode arises when the person being spoken to is monolingual (in this case in language A), and/or the topic, the situation, or the purpose of interaction require that only one language be spoken to the exclusion of the other(s). it is in this mode that interferences, that is, speaker-specific deviations from the language being spoken due to the influence of the other deactivated language, are most visible (they can also occur in a bilingual mode but are difficult to separate from other forms of language mixing such as code-switching or borrowing).

In position two, the speaker is in an intermediary mode. Language A is still the most active language, but language B is also partly activated. This kind of mode arises, for example, when a bilingual is speaking to another bilingual who does not wish to use the other language (in this case language B), or when a bilingual is interacting with a person who has limited knowledge of the other language.

In position three, the speaker is at the bilingual end of the continuum. Both languages are active but language B is slightly less active than language A as it is not the current language of communication. This is the kind of mode bilinguals find themselves in when they are interacting with other bilinguals who share their two languages and feel comfortable mixing languages. They usually first adopt a base language to use together (language A here) but the other language, often referred to as the guest language, is available in case it is needed in the form of code-switches and borrowings.

A code-switch is a complete shift to the other language for a word, a phrase, or a sentence, whereas a borrowing is a word or short expression taken from the less activated language and adapted morphosyntactically (and sometimes phonologically) into the base language (Grosjean, 1998: 137). Borrowings can involve both the form and the content of a word (called nonce borrowings) or simply the content (called loan shifts). Of course, a change in situation or topic may lead to a change of base language. Bilinguals differ among themselves as to the extent they travel along the continuum. Some rarely find themselves at the bilingual end whereas others rarely leave this end (bilingual communities).

Grosjean (2001: 2) points to the importance of taking into account language mode for better understanding of data from various bilingual populations. Each language mode will have an impact on language production (and perception) by the bilingual, and any conclusions drawn about such production (perception) can only be interpreted within the context of which language was more activated. It can partly account for problematic ambiguous findings relating to topics as language representation and processing, interference, code-switching, and language mixing in bilingual children, aphasics, etc. For instance, researchers who have examined bilingual language production have often reported instances of interference. The problem is that it is not always clear what is meant by this term (also called transfer or transference). For Weinreich (1953), interferences are instances if deviation from the norms of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language. Haugen (1956) refers to interference as the overlapping of two languages, Mackey talks of the use of features belonging to one language while speaking or writing another, and for Clyne (1972), transference is the adoption of any elements or features from the other language.

A direct result of this broad view is that interferences observed in linguistic studies often correspond to borrowings and even code-switches. As stated by Grosjean (2000: 13), 'we will never get to the bottom of this terminological problem, and we will never isolate interferences from code-switches and borrowings in bilingual speech, if we do not take into account (and control for) the language mode bilinguals and language learners are in when they are being studied'. In other words, language interference can only be identified correctly if deviations from the language being spoken due to influences from the other language took place in a near-monolingual mode, i.e. when the other language was almost deactivated. Grosjean (1982: 299) refers to interference as the involuntary influence of one language on the other which becomes quite apparent when a bilingual is speaking to a monolingual.

When interferences occur in a bilingual or intermediate mode, they are very difficult to separate from other forms of language mixing, especially borrowings. What might appear to be interference could be a guest element or structure produced by the speaker who is aware that his or her interlocutor can understand mixed language. Language mode is a variable to be studied independently (one will need to investigate ways of determining the bilingual's position on the continuum, among other things) but it is also a variable to control for. Failure to control for the language mode has important implications for the way in which findings are interpreted (Grosjean, 1998: 140). For instance, Schnitzer & Krasinski (1994: 616) give examples of interferences by their son Fernando that include dark [ɫ] for clear [l] in Spanish from age 3;3 to 3;4, and Spanish [r] for English [ɾ] (allophone of /t/) at age 3;8. These are considered errors in judgment as to which phonological system is appropriate; however, there was no mention of the language context in which they were produced.

Bilinguals who are strongly dominant in one language do more language mixing when speaking their weaker language than their stronger language, and may simply not be able to control language mode in the same way as less dominant or balanced bilinguals. Although they may deactivate their stronger language in a monolingual environment that only requires the weaker language, that language will simply not be developed enough to allow them to stay in a monolingual mode (Grosjean, 2001: 21).

Grosjean (2001: 8) notes that even though language mode has been alluded to by several researchers over the years, it has not been the object of systematic study until quite recently. A number of studies have found evidence for language mode in bilinguals (see studies reviewed by Grosjean, 1998; 2001). In many cases, the language mode continuum concept has offered a new approach to studying variable code-switching patterns within and between communities, because it can help predict the frequency and type of switching that takes place. For example, Lanza (1992) found that the same child mixed languages much more when in a bilingual context (represented by her father) than in a monolingual context (represented by her mother). Similarly, Treffers-Daller (1998) found different code-switching patterns by the same bilingual speaker depending on factors such as the interlocutor (familiar *versus* unfamiliar, monolingual *versus* bilingual), the context (e.g. formal *versus* informal) and the topic (e.g. questionnaire *versus* chat).

In speech perception, Elman, Diehl, & Buchwald (1977) tested Spanish-English bilinguals on voice onset time (VOT) continua (ba-pa; da-ta; ga-ka) and obtained identification curves in English and in Spanish. The language sets were obtained by changing experimenters (one English, one Spanish), the settings, and the language instructions. The bilinguals behaved like English listeners when in an English mode and like Spanish listeners when in a Spanish mode (i.e. they showed a perceptual boundary shift). More discussion on perception and production studies that have taken the language mode into account will be discussed in the relevant chapters. The current study will also offer further evidence that there are different phonological modes in the bilingual (regardless of whether or not any one mode is completely monolingual), and that each language mode has a different impact on the bilingual's language production.

1.7 General summary

In this chapter, I have approached the issue of whether the bilingual starts with one or two phonological systems by broadening the view of what a phonological system consists of. This was done by considering issues in monolingual phonological acquisition, including early perceptual abilities of the child, individual differences in development, and the role of input in the acquisition of phonology. With respect to the last issue, I have used insights from recent theories of speech perception and from studies on variability in speech to suggest that there are no simple targets for any child and that the child's job in gaining knowledge about sounds involves much more than the acquisition of contrasts. Studies of bilingual development, however, have mainly concentrated on a monolithic view of the two languages under investigation and have mainly taken the child's acquisition of contrastive sounds in either language as evidence for phonological differentiation. This study adopts a different stance on bilingual phonological acquisition in that (i) it investigates the relationship between the child's variable phonological input and the output; and (ii) examines socio-phonetic aspects of acquisition that are systematic but not necessarily crucial for signaling contrast. In the next chapter, I move on to the methodology.

CHAPTER TWO

The current study

2.0 Introduction

This study consists of three case-studies of English-Arabic bilinguals who were born and raised in Yorkshire, England. For each bilingual, an investigation was made of their family and friendship network in order to obtain as much information as possible about the sociolinguistic environment that is available for the bilinguals. Data were collected from the bilinguals, aged-matched controls from each language, and the parents of all bilingual and monolingual children. The study is cross-sectional in that three age groups were investigated and were taped while they were engaged in different activities over a short period of time. Furthermore, two of the bilingual subjects were taped twice with a time lapse of 18 months, which provided developmental data that were analysed separately (see Chapter Five). The study is experimental in nature, as it investigates the bilinguals' production of particular phonological variables of interest and uses quantitative data analysis to interpret the bilingual's linguistic behaviour.

In this chapter, the methodology that was designed for the study will be described in Section 2.1, followed for preliminary results from an accent rating experiment which was designed to obtain an overall impression of the bilingual's accent in English before looking at more detailed data in the following chapter (Section 2.2).

2.1 Methodology

The methodology for this study was developed following a pilot study (Khattab, 1998) which was conducted with two of the bilingual subjects who also took part in the current study. For the pilot study, two brothers from Leeds were taped while they named items in a picture-book designed to elicit a variety of sounds in English and Arabic. Analysis of the subjects' productions revealed interesting patterns which provided evidence that the children had developed two separate phonological systems for English and Arabic, though there were signs of possible interaction between the two languages. However, the interpretation of these patterns proved difficult in some respects, due to the lack of comparative material from monolingual children and adults from the same community. Some questions were therefore left unanswered, especially with regards to the fact that some 'atypical' patterns that were found could have been (i) the result of the children's bilingual background (including their parents' production); (ii) part of monolingual developmental patterns; and/or (iii) part of adult monolingual patterns, though not necessarily the ones described in the literature.

The pilot study enabled me to decide on (i) the type of control subjects (children and adults) that would be needed in order to be able to interpret the bilingual subjects' linguistic behaviour (Section 2.1.1); (ii) the kind of qualitative data that would be required to find out about the sociolinguistic background of the subjects (Section 2.1.2); (iii) the kind of data collection procedure that would be suitable for children and adults and that would yield a range of speech styles that vary from controlled to free and that cover both languages under investigation (Section 2.1.3.1 and 2.1.3.2); (iv) the linguistic variables that are relevant to the bilinguals' respective languages and communities and that would yield information on the bilinguals' communicative ability as well as their ability to develop and maintain a separate phonological system for each of their languages (Section 2.1.3.3).

2.1.1 The subjects

Three English-Arabic bilingual subjects were chosen for the current study. These were the two brothers from Leeds (aged 7 and 10), and also a 5-year-old girl from York. The choice of bilinguals was made depending on the availability of Lebanese families in the Yorkshire area, as it was necessary to control for the English and Arabic varieties being examined as much as possible. The three children all belong to families where both parents are Lebanese immigrants who have been living in the UK for a period of 10 to 15 years. Both families live in neighbourhoods where most of the residents are native English speakers from a variety of UK origins (see Section 2.1.2), and all three children go to mainstream English schools and interact mainly with native English speakers. For this reason, it was decided that including monolingual friends of the bilinguals and the parents of both groups would help establish the targets that the bilingual children are aiming for in acquiring English by offering valuable information on the type of speech patterns that are used by people that the bilinguals frequently interact with. Similarly, since not enough is known about speech developmental patterns in Arabic or about the phonetics and phonology of Lebanese Arabic, three monolingual Lebanese children and their parents were included in the study; they were chosen from the same district area in the Lebanon (West Beirut) as the bilingual parents.

A total number of 23 subjects were recorded, as shown in Table 2.1. With respect to the children, there are three age groups (5, 7, and 10), each consisting of one bilingual and one monolingual subject from each language. All the subjects in a given group are of the same sex (the five-year-olds are females whereas the seven- and ten-year-olds are males). Two of the monolingual Arabic children and two of the bilingual children are siblings. For this reason, there are only 4 parents in each of the bilingual and monolingual Arabic groups while there are 6 parents in the monolingual English group. In Section 2.1.2, a detailed profile will be given of each of the subjects.

Table 2.1: Subjects, listed in numbers and grouped according to age.

Language groups	**Age 5 (F)**	**Age 7 (M)**	**Age 10 (M)**	**Adults**
Monolingual English (E)	1	1	1	6
Bilingual (B)	1	1	1	4
Monolingual Arabic (A)	1	1	1	4
Total: 23	**3**	**3**	**3**	**14**

2.1.2 Sociolinguistic background of the subjects

After the two UK Lebanese families were located, several visits were made to the bilingual subjects' homes before the start of the study in order to get to know the children and their families. Then a series of sociolinguistic questionnaires and interviews were conducted with the parents, their children, and the children's school teachers in order to gather as much information as possible about the bilingual subjects' language and social background (Appendices 1, 2, and 3). Issues such as the language(s) spoken initially by each child, the language(s) used at home and with friends, the attitudes of both parents and children towards each of the languages and respective communities, and future plans to move back to the Lebanon were investigated.

Monolingual English children were then chosen for the study among friends that the bilinguals stated spending time with the most and with whom they had been friends since nursery or playgroup. Visits were made to the monolingual subjects' homes in order to meet the families and conduct interviews and recordings with the children and their parents. As for the monolingual Arabic children and their families, these were interviewed and recorded in their homes in the Lebanon.

Table 2.2 summarises the backgrounds of the 23 subjects who were taped for this study. For the purpose of brevity, initials will be used for the subjects in the presentation of results in Chapters Three to Five. 'E', 'B', and 'A' stand for the three language groups (English, bilingual, and Arabic respectively). Numbers have been added to the children's initials to represent their age, e.g. E5; B7; A10, etc., and the adults were divided into males and females. Therefore each adult will have two initials, one for the language group and the other for gender, along with one of the numbers 5, 7, or 10 to help identify the child/children of each adult, e.g. EM5; BF7; AM10, etc.

Table 2.2: Details of the 23 subjects who were taped for this study.

Child					Parents			
Age group 5		**Age**	**Sex**	**Origin**	**Mother**		**Father**	
Monolingual E	E5	5;5	F	York	EF5	Kennick	EM5	Leicester
Bilingual	B5	5;6	F	York	BF5	Beirut	BM5	Beirut
Monolingual A	A5	5;4	F	Beirut	AF5	Beirut	AM5	Beirut

Child					Parents			
Age group 7		**Age**	**Sex**	**Origin**	**Mother**		**Father**	
Monolingual E	E7	7;5	M	Leeds	EF7	Stockton	EM7	Norwich
Bilingual	B7	7;1	M	Leeds	BF7	Beirut	BM7	Beirut
Monolingual A	A7	7;4	M	Beirut	AF5	Beirut	AM5	Beirut

Child					Parents			
Age group10		**Age**	**Sex**	**Origin**	**Mother**		**Father**	
Monolingual E	E10	10;3	M	Manchester	EF10	York	EM10	London
Bilingual	B10	10;2	M	Leeds	BF7	Beirut	BM7	Beirut
Monolingual A	A10	10;3	M	Beirut	AF10	Beirut	AM10	Beirut

2.1.2.1 The monolingual English subjects

The monolingual English subjects are close friends of the bilinguals who live in the same area. Although all three monolinguals were born and raised in Yorkshire, their parents come from different areas in the UK and have lived in several places before moving to Yorkshire. Below is a detailed report on each of the monolingual English children's background in increasing age order.

Lissa (E5)

E5's mother, EF5, was born in Kennick, Staffordshire, but grew up in Northumberland until she was 18. She moved to Newcastle for university education, and then moved to York for work where she had been living for 16 years at the time of the interview. E5's father, EM5, was born in Kibworth (South Leicester), and grew up there until he was 18. He then moved to York for university education and work and had been living there for 17 years at the time of the interview. The couple live in a middle class neighbourhood inhabited by families who have been living there for a long time and who were reported as being mainly from Yorkshire. Their daughter, E5, was born and raised in York, and has a three-year-old brother.

When asked to comment on her own accent, EF5 noted that it sounded more southern than it actually was, and that people often assumed she was from the South when in fact she grew up in Northumberland. She later admitted that her pronunciation of the same words was often variable from one production to the other, and gave the example of 'castle' which she reported as producing sometimes as [ˈkʰasɫ̩] and other times as [ˈkʰɑːsɫ̩] (similar variable pronunciations were noticed during the interview).

EM5, on the other hand, maintained that he had a Leicestershire accent, giving illustrations of his productions of 'look' and 'book' as [lu:k] and [bu:k], but admitted that he had acquired certain Yorkshire 'words'.

When asked to comment on E5's accent, EF5 thought that her daughter had an accent that was similar to her own, but that had more Yorkshire influence due to the school and neighbourhood. She reported variable pronunciation in her daughter's production, and noted that E5's friends and school environment consist of children from a mixed regional background. However, the child had also been greatly influenced by input from one of her teachers, who was described as 'very Yorkshire'. EM5, on the other hand, noted that his daughter's accent 'does not sound much like an accent', but that it was not dissimilar to that of her parents. He reiterated his wife's remark on the availability of a mixture of English accents that E5 is exposed to from the school, mentioning friends that she plays with whose families have only been living in York for a couple of years. Interestingly, EM5 also noted that this daughter often came up with a lot of American words and he 'did not know where she got them from', but maintained that she also produced a lot of Yorkshire words.

In summary, both parents consider that E5's accent has equal influence from their own accents and from that of the school and neighbourhood.

William (E7)

E7's mother, EF7, was born in Stockton-on-Tees in the North East of England and grew up there until she was 18, apart from one year that she spent in Denver (USA) when she was 13. At 18, EF7 moved to Birmingham for a university degree and stayed there for four years. She then lived in each of Milton Keynes (four years) and Wakefield (three years) before moving to Leeds when she was 29, and had been living there for 13 years at the time of the interview. E7's father, EM7, was born in Norwich, but only lived there till he was three when his family moved to Yorkshire and lived near York. He then went to Leeds University when he was 18 and had been living in Leeds for 29 years at the time of the interview. The couple live in a middle class neighbourhood inhabited by families who they reported as not having lived there for a long time and who came from various areas of the UK. Their son, E7, was born and raised in Leeds, and has a 10-year-old sister who was also brought up in Leeds.

When asked to comment on her own accent, EF7 described it as 'neutral' and said that she normally 'picks up' some aspects of the local accent wherever she lives. She illustrated that by saying that when she lived in Milton Keynes, she used to say [pʰɑ:θ] and [bɑ:θ] for 'path' and 'bath', whereas when she moved back to the north she started saying [pʰaθ] and [baθ] again. Similarly, EF7 noticed that after she came back from a year in the US, her friends thought that she had developed an American accent. Surprisingly, EM7 also described his accent as 'neutral' although he is one of only two subjects in this study who has lived in Yorkshire all his life. He portrayed his accent as 'the sort of accent you get from having moved a lot and been to university, where you are influenced from lots of people'.

When asked to comment on E7's accent, EF7 said that she would not describe it as a strong Yorkshire accent, but that people from the south of England would definitely identify him as being from the north. She noted that his accent was 'quite subtle and light', 'a bit more Leeds than her own', and 'similar to that of other children in the neighbourhood whose parents come from outside Leeds but the children have a little bit of the local accent'. EF7 further attributed the 'subtle' accent of her son to the fact that they live in an area full of 'professionals who move around the country a lot'. She noted that E7 is exposed to a much wider range of accents than he would if they were living in another area of Leeds (probably referring to a working class area). EF7's observation is interesting, as it reveals a lot about the sociolinguistic situation in the community being studied: the subjects belong to middle class families, the first generation of speakers (the parents in this study) are mainly outsiders, and the 2nd generation (the children in this study) might be contributing to the phenomenon of 'accent leveling'. Accent levelling is a process whereby differences between regional varieties are reduced, features which make varieties distinctive disappear, and new features emerge and are adopted by speakers over a wide geographical area (Williams & Kerswill, 1999: 149). EM7 reiterated his wife's opinion about his son's accent being a bit more like that of his school friends than those of his parents, but maintained that he would not describe it as Yorkshire due to the kind of area they live in.

Andrew (E10)

E10's mother, EF10 was born in York and lived there till she was 22. Having completed a degree in nursing, EF10 then moved to Leeds for work for five years before she married and moved to London for a few months. She then lived in each of Surrey (four years), Baltimore, US, (18 months), and Manchester (five years) before the couple moved back to Leeds and had been living there for six years at the time of the interview. E10's father, EM10, is of Polish extraction, though he was born in London and grew up there till he was 18. EM10 then moved to Leeds for education and lived there for 10 years. After marrying EF10, EM10 then lived in the same places as mentioned above for EF10. EM10 considers himself as being bilingual in English and Polish, but has not taught his son any Polish, and, as a consequence of that, E10 grew up as a monolingual English speaker. E10 was born in Manchester and lived there until he was five, when the family moved to Leeds. E10 has two older siblings (aged 13 and 15), and a younger sister (aged 9). Like the other two families, E10's family lives in a middle class area.

When asked to describe her accent, EF10 is the only subject from this study who considered herself as having a Yorkshire accent, and underlined the fact that her accent was North rather than West Yorkshire, as she thought the two accents differed considerably. EM10 initially called his accent 'middle class' and, when asked to elaborate, admitted that it depends on where he is: 'when I go home I speak with a Cockney accent to my friends, and when I'm here I use northern English pronunciation with slightly truncated vowels; if anything, my accent is a little bit of a chameleon'.

When asked to comment on E10's accent, EF10 described it as 'a bit of a hybrid' due to the fact that his father is from London but that the son had acquired 'northern sounds' as well. EF10 further said that she could not 'place' her son's accent and would not say he had a Leeds accent, despite the fact that most of his friends are originally from Leeds and that he mixes with children from her extended family, who come from York, Hull, and other nearby areas. EM10, on the other hand, described his son's accent as 'posh Leeds', and as being 'not as strong as that of his friends whose parents were born and brought up in Leeds'. EM10 later admitted that his son only sounds like he has a Yorkshire accent when the family visits relatives in London.

Additional data

The profile of the monolingual families in this study is representative of that of many families living in cities like Leeds and York, where extensive geographical mobility has led to a great number of people from different regions in the UK moving to urban areas in search for work and/or education. The children of these families are therefore likely to be exposed to a different accent or accents at home than the one(s) available in their community. Due to the fact that only 2 of the parents from this study were actually brought up in Yorkshire (EM7 and EF10), Grabe & Nolan's (2001) speech data from Leeds were also used for the study. Their speakers had lived in Leeds all their lives and are therefore more likely to produce patterns that are typical of the Leeds accent than any of the Leeds subjects in this study. Although the Leeds IViE speakers all come from a working class background, the availability of Grabe & Nolan's recordings of these speakers was very helpful in that it offered data on present-day Leeds speech, even if from a small sample of speakers. The data were collected from twelve 16-year-olds at urban secondary schools, and the speech of ten of the speakers was analysed for this study. No comparable data were found for York.

2.1.2.2 The monolingual Arabic subjects

The monolingual[1] Arabic families were chosen from the same district as the bilinguals' parents in the Lebanon. This arrangement was made in order to control for dialectal differences as much as possible. Among the children, A5 and A7 are brother and sister. The children and their parents were born in Beirut and had been living there all their lives at the time of the interview. Like many families in the Lebanon, the children are exposed to French in schools from the age of three. As a result of this, many French words have infiltrated the vocabulary of the children, but the main language used at home and in the community (for these subjects at least) is still Arabic. The children have also come into contact with English from the media (mainly television and radio), but during the interviews all children stated that they had no knowledge of English. As for the parents, they all speak French which they learned as a foreign language at school, and one of them speaks English as a third language (AM5).

2.1.2.3 The bilingual subjects

The bilingual subjects are children of two Lebanese families who live in York and Leeds respectively. The social circumstances in which the families moved to the UK are quite similar. The fathers came over first for higher degrees, then stayed to work and brought their wives over to start a family. Below is a detailed report on each of the bilingual children's background in increasing age order.

Maguy (B5)

B5's mother, BF5, was born in Beirut and grew up there till she was 19. She then got married and moved to York, and had been living there for nine years at the time of the interview. BF5 speaks French as a second language (from the age of three) and English as third language (from the age of 11). She teaches French in a school in York, and describes herself as an intermediate speaker of English. B5's father, BM5, was born in Beirut and grew up there till he was 18, and had the same foreign language background as BF5. He then moved to London for a Masters degree and stayed for a PhD. After his degree, he worked as a researcher at his university for a couple of years, then went back to the Lebanon to get married. He brought his wife to York, where he was offered an engineering job at the University of York. At the time of the interview, BM5 had been living in the UK for 15 years and described himself as an advanced speaker of English.

---

[1] Note that although the Lebanese-based families are labeled 'monolingual', they will have been exposed to French and/or English as a foreign language in schools, universities, etc.

B5 was born and brought up in York, and has a 2-year-old brother. B5's first words were in English, soon after she started going to nursery at the age of 1;0. B5 had English friends and spoke English exclusively until the age of 2;5, when she was taken to the Lebanon for the first time and spent a period of 6 months surrounded by Arabic-speaking friends and relatives. By the end of her stay, B5 was reported as speaking Arabic as fluently as any child her age, but her Arabic use soon decreased after she came back to the UK. Since then the family has been back to the Lebanon once for a shorter visit when B5 was 4;0, and her Arabic use was again noted to have increased by the end of her stay. But while in the UK B5 mainly has English friends from the neighbourhood and from school (the same neighbourhood and school as E5), and the input she receives is predominantly English, including sources of entertainment such TV, music, games, and reading. B5 also speaks English to her parents and to her younger brother. B5 has just started reading in English, but does not read or write in Arabic. No extra effort was being made to teach B5 reading and writing in Arabic, and all the bedtime reading activities with either parent were in English.

Both parents consider English to be B5's dominant language. Questionnaire data (Appendix one) revealed that her mother rated her English at 5 out of 5, and her Arabic at 4 out of 5, while her father gave her 5 out of 5 for both, judging her Arabic performance following a trip to the Lebanon. B5's school teacher also rated her English at 5 out of 5, and said that she definitely sounded native-like in English, had very good interpersonal skills, and a healthy social life with her classmates. When asked to comment about her accent, the teacher said that her accent would be a bit difficult to identify, as was the case with other classmates, since only a few of them had a Yorkshire accent, while a lot of the students 'had no regional accent'. The teacher added that while three quarters of the class consisted of native English speakers, the rest were from a variety of language backgrounds.

When asked about the language(s) used in the house, BF5 reported using both Arabic and English regularly with her husband and children, but later admitted that she used more Arabic when speaking to her husband, and more English when speaking to her children. BM5 reported the same, though he added that while he mainly spoke English to B5 while she was growing up, he now tries to speak Arabic to her 2-year-old brother after he noticed that B5's Arabic suffered a delay because of the little input she was receiving. BM5 actually noted that a lot of Arabic was spoken in the house while B5 was growing up, but little of it was being directed at her and therefore she did not benefit much from this input.

The children's questionnaire (Appendix one), which was conducted in English, did not prove very successful with B5, partly due to her young age and therefore the possibility that she found some of the questions difficult. I also had the impression she often gave me the answers she thought I wanted to hear. For instance, her answers to questions like which language she knew better, which language she preferred, and which language she used with her parents was always 'Arabic', though during the Arabic sessions with her mother she often resisted speaking Arabic and resorted to English instead. Similarly, though B5 said that she liked living in the UK and had lots of English friends, she later said she preferred living in Beirut (her mother said that B5 was normally spoiled with presents and food from the relatives every time the family went back to the Lebanon, and therefore associated their trips with these things).

B5's family reported being generally happy with their lives in the UK and content with having mainly English friends. There are no Arab families that they know in York, and the only other Arabic acquaintances in the UK include B7 and B10's family in Leeds, and another family in Sheffield. The three families get together on a monthly basis, but note that the children speak English to one another most of the time. Both BF5 and BM5 said they had no immediate plans to go back to the Lebanon. The couple were also happy with the fact that B5 was growing up bilingual, and did not mind that English was her stronger language. In fact, her father saw that as an advantage for the future in terms of B5's career opportunities (whether in the UK or in the Lebanon), and both parents seemed to be content with the fact B5 spoke enough Arabic to communicate with her Arabic-speaking relatives.

Mazen (B7) and Mohammed (B10)

B7 and B10 are brothers from Leeds. For practical reason, the initials used for their parents will be BF7 and BM7. BF7 was born in Beirut, and grew up there until she was 20, when she married BM7 and moved to the UK. Like BF5, BF7 speaks French as a second language (from the age of three) and English as a third language (from the age of 11). When she arrived in the UK, BF7 attended an English course at a college and mainly stayed at home with her husband who was doing his PhD in Sheffield. When the family moved to Leeds two years later, BF7 went on to work as an administrative assistant in the school that her children are now attending. At the time of the interview, BF7 had been living in the UK for 12 years and described herself as an intermediate speaker of English.

BM7 was born in Beirut and lived there until he was 27, having had a similar foreign language background to the rest of the three bilinguals' parents. He then moved to Edinburgh for a Masters degree, and after that started a PhD at Sheffield. Towards the end of the degree, BM7 married BF7, and the couple lived in Sheffield for two years before moving to Leeds where BM7 now works as an engineer. At the time of the interview, BM7 had been living in the UK for 16 years, and described himself as being an intermediate speaker of English.

In both Sheffield and Leeds, the family lived in English-speaking neighbourhoods and mainly had English-speaking friends. However, while the friends they knew in Sheffield were mainly native-speakers of British English, in Leeds, the couple reported making friends with families from different language backgrounds including Pakistani, Iraqi, and Russian, most of whom were work acquaintances. Both BF7 and BM7 reported speaking only Arabic to their children, as they were aware that they were the only source of Arabic to the children. The parents were extremely keen for the children to acquire not only speaking, but also reading and writing in Arabic. Both parents later admitted that they would sometimes use English 'words' when they suspected that the children did not understand what they said in Arabic, or when they were using lexical items that they are in the habit of producing in English. Informal observation of the family's interactions revealed a lot of code-switching, though it seemed to be below the level of consciousness of both parents and children. Still, the Leeds couple certainly spoke more Arabic between them and to their children than the York couple.

BF7 and BM7's ultimate aim is to return to the Lebanon in the future, so they want their children to be proficient enough in Arabic to be able to lead a successful life there. They are also keen on teaching them literacy for religious reasons (so that the children can learn to read the Koran). For these reasons, at the time of the recordings for the current study, B7 and B10 had been attending a weekend Arabic school in Leeds for 18 months, where around 80 children from different Arab countries (but not the Lebanon) learned Arabic literacy. Apart from the Arabic school, the parents also expose their children to Arabic TV channels through digital media, along with Arabic music and books which they take turns in reading to them.

B10 was born in Sheffield, where he attended nursery three days a week between the ages of 1;2 and 1;5. When the family moved to Leeds, he was looked after by an English child minder for three months, but later the family went to the Lebanon for a period of six months. In the Lebanon, B10 was mainly exposed to Arabic and, on the family's return to the UK, the mother noticed a considerable increase in his understanding of Arabic and in his Arabic vocabulary. B10 then joined a full-time nursery at the age of 2;5, where most of his friends were native English speakers. At the age of 3;5, B10 had another extended exposure to Arabic in the Lebanon for a period of eight months, and his parents noticed more improvement in his Arabic skills. However, since the family's return to the UK, B10 has mainly had English-speaking friends and has not been back to the Lebanon for more than one month every year.

B10's first words to his mother at 1;3 were Arabic, but after three months of nursery he showed signs of speaking and certainly understanding more English than Arabic. Between 1;5 and 2;5, B10 spoke mainly Arabic to his mother and English to his father. He thought his mother did not speak any English, and was very surprised when he heard her speak English to his nursery teacher when he was 2;5. Since then, he has tried speaking English to both of his parents, but they always resist answering him in English and encourage him to speak Arabic instead. At the time of the recording, B10 was reported as frequently code-switching with his parents and as mainly speaking English with his brother. However, they had both started producing more Arabic utterances since they started going to the weekend Arabic school.

B7 was born in Leeds, where he attended full-time nursery from the age of 1;0. The family followed the same policy of speaking Arabic to B7, while B10 often spoke English to him, and, only after encouragement from the family, made the occasional effort to speak Arabic to him. As opposed to B10, B7 never spent more than 3 months at a time in the Lebanon, and his pronunciation of Arabic has always been described as 'broken' and as being 'stranger' than that of B10 by parents and relatives. B7's first words were Arabic, but English soon took over due to the nature of the environment. During visits to the Lebanon, B7 frequently asks his mother for translations of English words so that he can communicate with relatives, whereas B10 is becoming less reliant on his parents for help.

Both parents consider English to be B7 and B10's dominant language, but think that B7 is slightly better at English than B10, while B10 is slightly better at Arabic than B7. Still, BF7 rated the two children's English at 5 out of 5, and B7's Arabic at 3 out of 5, and that of B10 at 4 out of 5. The father rated both the English and the Arabic of the two children at 4 out of 5, but maintained that English was dominant, especially when they played together. The parents noticed that the weekend Arabic school had certainly improved the two children's pronunciation as well as literacy skills, but not their language preference.

B7 and B10's school teacher had taught both of them since they joined the school, and also rated their English at 5 out of 5. He said that they both definitely sounded native-like in English, but that they were both shy when they started attending the school and lacked confidence. In the early stages, B7 received extra help to improve his literacy skills (spelling and word analysis skills), and B10 received extra help to improve his reading and grammar skills, but the teacher did not attribute any of the difficulties that the brothers faced to their bilingual background, and instead maintained that their skills developed similarly to those of students in general.

When asked to comment on the children's accent, the teacher thought the two brothers sounded 'more polite' than their English-speaking friends from Leeds some of whom sounded 'broad' and 'sometimes ate half their words', but later contradicted himself and noted that most of their classmates had a 'middle class' accent that was not a typical Leeds accent, and that 30% of the students came from other language backgrounds. The teacher mentioned an English-Panjabi bilingual student who was a close friend of B10 until the previous year and who he thought had a great influence on B10's English accent. This observation turned out to be relevant during the interpretation of some of B10's results (see Section 2.4). Apart from that, the teacher noted that the brothers had a healthy social life in the school, were well-adjusted, and were both keen on doing well.

B7 and B10's parents are also satisfied with their social life in the UK, but, unlike B5's family, often feel homesick and are keen to return to the Lebanon in the future. Though there are no other Lebanese families that the couple know in Leeds, they often actively seek Arabic-speaking friends and prefer socialising with them to English-speaking families mainly due to the reported ease of communication in Arabic. The couple are, however, very positive about their children's growing up bilingual and consider that as a great advantage for their future, whether in the UK or the Lebanon.

The children stated being happy with their lives and friends in the UK, but looked forwards to their visits to the Lebanon where, like B5, they are spoiled by family and friends. Both children reported English to be their preferred and dominant language, and said they preferred the English school to the Arabic one, mainly because the Arabic school takes up part of their weekend, but also because they are not familiar with the teachers' Iraqi accent and often struggle to understand what the teachers are saying. The children still expressed their interest in learning Arabic, especially due to the fact they receive huge encouragement from the parents, but also because they can 'show off their skills' to their Arab relatives.

2.1.2.4 Summary

The two Lebanese families live in mainly English-speaking neighbourhoods and the children attend mainstream English schools, so the only contact that the children have with Arabic is from their parents and a couple of Lebanese families living in other cities and that they only occasionally get in contact with. All the parents are native speakers of Arabic and, while B5's parents use both English and Arabic at home, B7 and B10's parents use mainly Arabic, but also use English in public places or in the presence of monolingual English speakers. Code-switching between the two languages is a common feature in the speech of parents and children alike in both families. B5 was exposed to both English and Arabic from birth, while B7 and B10 were exposed to Arabic first, but all three children started attending English play groups and nurseries at an early age and all three children are English-dominant. Both families are keen on bringing up the children to be bilinguals because they have positive attitudes towards the two languages for a multitude of reasons (social life and work in the UK on the one hand, and religion and visits to the Lebanon on the other).

2.1.3 Procedure

2.1.3.1 Data collection

After having gathered qualitative data from the interviews and questionnaires with the children, their parents, and the bilingual children's school teachers, speech data elicitation took place in each of the subjects' homes. The decision to record the subjects in their homes was an attempt to provide a relaxing and familiar atmosphere for both children and adults despite the presence of a tape recorder, and to avoid the formal laboratory atmosphere which might intimidate the children or encourage the adults to adopt a formal style of speech. Tape-recording sessions were designed around four types of activities.

First, picture-naming games were used for all nine children in the study in order to elicit English and Arabic target words in isolation and short utterances. Around 200 words that were taken from age-appropriate vocabulary lists and that can be represented in pictures were compiled in the form of a picture-book (Appendix two). The adults were given reading lists containing the same target words that the children produced, and were instructed to produce these words as they normally would in their everyday life (each adult was given time to go through the list and familiarise themselves with it). Such instruction was particularly important for the elicitation of Lebanese Arabic, due to the considerable difference between written and spoken Arabic, the aim being for the speakers to produce words in their dialect. Written Arabic represents Classical Arabic, which is used in religion and liturgical matter (Hussein, 1980: 82), and Modern Standard Arabic, which is used in mass media such as news reading and in inter-dialectal situations (during communications between people who speak two different Arabic dialects). Colloquial Arabic, on the other hand, represents spoken Arabic in everyday situations between members of a given community (Mitchell, 1993).

Second, story-telling activities were used for both children and adults in order to obtain running speech that is minimally controlled. Mayer's (1969) picture stories *Frog, where are you?* and *One frog too many* were used for the elicitation of narratives. The books consist of 24 pictures each without accompanying words, and have been used by other researchers on bilingual and monolingual language acquisition (Berman and Slobin, 1994). It was therefore thought that using these books for the study would allow other researchers to use the material for other linguistic types of analysis and would allow cross-linguistic comparisons of findings in the future.

While each of the monolingual children and adults were recorded once, the bilingual children and their parents were recorded twice, following a one-language-per-session approach (Grosjean, 1998) as an attempt to maximise the activation of only one of the two languages each time (see discussion on the bilingual's language modes, Chapter One). While I conducted the sessions with the children in English, the mothers were asked to conduct the Arabic sessions on the basis that the children would be more likely to use Arabic with their parents than with anybody else in their environment.

There were two problems with this approach. First, while the children used only English in the English sessions, they frequently reverted to code-switching during the Arabic sessions or responded in English even when the mothers were asking questions in Arabic. Such behaviour was to be expected considering the fact that the bilinguals in this study are English dominant and that, while many English speakers in their environment are monolingual, all the Arabic speakers they know are either bilingual or speak English as a foreign language. However, the code-switched utterances proved to be a rich source of comparative phonetic data, and were therefore analysed and interpreted separately from the single-language utterances. Their inclusion proved significant in the overall interpretation of the bilinguals' data.

Second, even though the children produced only English during the sessions with me, they knew that I spoke Arabic and therefore the sessions did not necessarily elicit the kind of linguistic behaviour that the bilinguals would normally exhibit when communicating with monolingual English speakers. According to Grosjean (1998) the bilingual's two languages might still be highly activated when (s)he is communicating with another bilingual, even if only one language is being used.

For this reason, a third set of data was collected, consisting of free-play sessions during which each of the bilingual children was paired with their monolingual friend of the same age and left alone in the room with games for around 45 minutes. A similar session was set up by pairing two of the bilinguals together, B7 and B10, in order to compare their linguistic behaviour during a bilingual-monolingual interaction and a bilingual-bilingual interaction. The free-play sessions provided data that were near-naturalistic, as the tape recorder was hidden in a corner of the room. After a while, the children seemed to forget the presence of the small microphones, especially as I was not in the room. The data from these sessions therefore played an important role in providing support for the results obtained from the picture-naming and story-telling activities, and for comparing the bilinguals' performance on a range of linguistic activities and interlocutors.

Finally, a fourth and final type of data used for the study consisted of the interviews with the adults, who were all tape-recorded while being asked to discuss biographical data about themselves and their children. It was thought that, although the adults would always be more aware of the presence of a tape-recorder than the children, the interview sessions might still contain more naturalistic data than the reading lists and the story-telling activities. Appendix three shows a summary of the number and type of recordings that were conducted for the current study.

2.1.3.2 Data recording and analysis

A Tascam DA-P1 DAT recorder was used during all sessions, with Trantec external microphones clipped to the subjects' clothes, normally the collar or top pocket of a shirt. The quality of the recordings allowed both auditory and instrumental analysis of the data despite the fact that the sessions were not conducted in a laboratory. Common problems included noise due to the children moving around and touching their microphones, or due to background noise from other rooms in the house or outside. Still, the outcome was a compromise between high-quality recordings and naturalistic settings. Auditory analysis was conducted on all the data collected for the current study along with the Leeds corpus data, while instrumental analysis was only conducted on the words produced in isolation from the picture-naming activities for the children and the reading lists for the adults. Acoustic analysis allows minute differences in individual pronunciation to be measured, at a level of detail which is too fine to be perceptible to the best trained ears.

2.1.3.3 Linguistic variables

Each of the linguistic variables that were analysed will be discussed in detail in the relevant chapters. For this reason the variables will only be briefly introduced here.

/l/ production

/l/ was chosen due to the existence of different patterns for clear and dark variants in its production in English and Arabic that vary according to contextual and dialectal factors (e.g. Carter, 1999; Sproat & Fujimura, 1993; Shaheen, 1979). In Yorkshire, initial /l/'s are generally described as 'dark-ish' (Wells, 1982: 371), to denote an intermediate kind of /l/ in all environments. Apart from the clear and dark allophones, some dialects of English permit vocalisation of syllable-final /l/, and a range of back vocoids may be used as the reflex of [ɫ] e.g. 'milk' [miʊk], 'fill' [fɪɤ]. In Arabic, /l/ is clear in all word positions, with the exception of emphatic environments. The present study is based on an auditory and acoustic analysis of English and Lebanese onset and coda /l/'s as produced by the bilingual children and their parents on the one hand, and the monolingual English and Arabic children and parents on the other. The aim is to examine the extent to which children exposed to two languages acquire separate sociolinguistically appropriate production patterns for /l/ for each of their languages.

/r/ production

In most English accents, /r/ is produced as a voiced alveolar or post-alveolar approximant [ɹ] (Hughes & Trudgill, 1996: 90; Wells, 1982: 368). Moreover, for historical reasons, post-vocalic /r/ is absent before a consonant or in absolute final position in several accents of English (*farm* [fɑːm]; *far* [fɑː]) (Cruttenden, 1994: 268; Hughes & Trudgill, 1997: 60; Wells, 1982: 218). Arabic /r/, on the other hand, is normally a tap or a trill, depending on free and allophonic variation (e.g. Anani, 1985: 132; Shaheen, 1979: 142). Moreover, Arabic /r/ is produced in all pre- and post-vocalic contexts. The present study is based on an auditory and acoustic analysis of English and Lebanese onset and coda /r/'s as produced by the bilingual children and their parents on the one hand, and the monolingual English and Arabic children and parents on the other. The aim is to examine the extent to which children exposed to two languages acquire separate production patterns for /r/ for each of their languages.

Voice Onset Time (VOT)

English and Arabic vary considerably in their phonetic realisation of the stop voicing contrast. In English, utterance-initial VOICED stops are normally produced with short voicing lag or voicing lead, whereas VOICELESS stops are produced with long lag (Lisker & Abramson, 1971: 767; Weismer, 1980: 428). In Arabic, the contrast is often described as that of long lead for VOICED stops and short lag for VOICELESS stops (Flege & Port, 1981: 126; Yeni-Komshian, Caramazza & Preston, 1977: 35). The present study is based on a spectrographic analysis of English and Lebanese word-initial stops as produced by the bilingual subjects and their parents, along with the monolingual children and adults of either language. The aim is to examine the extent to which children exposed to two languages establish phonetically distinct contrasts for either language.

2.1.3.4 General aims of the study

The study is designed around five main questions, four of which will be repeated in Chapters Three to Five and will be answered in relation to the variable under study. These questions are:

1 Do English-Arabic bilinguals acquire separate sociolinguistically appropriate production patterns (for the variables under study) for each of their languages?

2 Are their patterns of production in each language similar to those of the monolingual controls in the study?

3 Are the patterns for the monolingual subjects in this study similar to the ones normally described in the literature and therefore expected for each language?

4 Are there signs of influence from one language to the other in the bilinguals' production? If so what are the factors that affect such influence and how are they related to the bilinguals' language modes?

5 What do the data tell us about the bilingual's processing of the two languages? This question is a very broad one and has many aspects to it. I am interested here in how the bilinguals learn, store, and use their two languages. I am also interested in the role of interaction, which has dominated the literature on bilingualism. An attempt to answer this question will therefore be made in the general discussion in Chapter Six.

2.2 Accent rating experiment

Before moving on to the detailed analysis of the subjects' production (Chapters 3-6), an overall impression of the bilingual children's accent in English was sought by running an accent rating experiment in order to find out whether listeners perceived the bilinguals' accent as native or non-native, and whether the accent could be traced to a particular regional origin. As there are only three bilingual children, a decision was made to include a selection of other subjects from the study, including the monolingual children and some of their parents, as well as some of the bilinguals' parents. The aim was to obtain a variety of English accents ranging from definitely non-native (the bilinguals' parents) to definitely native (the monolingual subjects) and to find out where native-English listeners would place the bilinguals. The task was made possible due to the existence of comparable data from the story-telling activities, since all the subjects used the same picture-books.

Speech files of 30-second duration from the story-telling data were chosen for 12 of the 23 subjects from the study, including the three bilingual children, the three monolingual English children, two of the bilinguals' parents, and three monolingual English parents, one of whom was included twice in order to test the reliability of the listeners' answers. The speech files were taken from different stages of the story, but never from the beginning, due to the fact that some speakers were a bit hesitant at first and took time to get into the narrative mood. The files were put in a random order before they were played to listeners (see below).

Thirty second and third year linguistics students at the University of York, all native speakers of English from different UK origins, took part in the experiment. The listeners were informed that they would be listening to short extracts from 12 speakers of different ages telling a story and that, after each of the recordings, they would be asked three questions about each speaker (Appendix four). The questions were:

a) Give each speaker a number on a scale from 1 to 4 ranging from:
   1 = definitely native   2 = probably native
   3 = probably non-native   4 = definitely non-native
b) If possible, explain your choice by referring to specific <u>linguistic features</u> (pronunciation, vocabulary use, etc.).
c) If possible, try to define each speaker's <u>accent</u> as narrowly as you can (e.g. Northern/Southern; Yorkshire; Lancashire; Manchester; Leeds; etc).

The sound files were played on a loud speaker in a lecture room, and the listeners were sat on individual desks and were given one minute after each recording to answer the questions by filling in cells in a table (Appendix four). Figure 2.1 and Table 2.3 shows the numerical results for the accent rating. In Figure 2.1, categories 1 and 2 ('definitely native' and 'probably native') have been combined, as have categories 3 and 4 ('probably non-native' and 'definitely non-native'), in order to initially compare 'native' and 'non-native' answers. In Table 2.3, the detailed ratings in numbers and percentages are given for each of the 12 subjects.

As expected, all of the listeners identified the monolingual English parents as native, and all but one of the listeners identified the bilinguals' parents as non-native. As for the children, despite the fact that the majority of choices for both monolingual and bilingual groups were in the 'native' categories, detailed results (Table 2.4) show more hesitation about choosing the 'definitely native' category, as many of the choices were in the 'probably native' categories. There are many reasons for this outcome.

With respect to the monolingual children, the younger children (E5 and E7) received more 'probably native' than E10 due to the fact that some listeners picked on developmental syntactic errors (e.g. 'he go') and overgeneralizations (e.g. 'bited') in the speech of E5, and described E7's speech as slow or hesitant (which it was). Although other listeners mentioned that they knew these features were part of native developmental patterns, they still chose category 2 rather than 1. For these same reasons, one listener thought E5 was 'probably non-native', and 4 other listeners thought the same for E7. E10, on the other hand, received no non-native choices and had the highest percentage of 'definitely native' choices (80%, as opposed to 50% for E7 and 40% for E5).

With respect to the bilingual children, the increase in 'definitely native' was, by contrast, inversely related to age and was more of a reflection of the actual accents and linguistic backgrounds of the three subjects. B5 received the highest percentage of 'definitely native' choices (80%, as opposed to 50% for both B7 and B10), and no choices from either of the 'non-native' categories. B7 and B10, on the other hand, received more 'probably native' choices than B5 due to observations such as 'not very fluent', 'some common syntactic errors but also some uncommon ones', and production of [t]'s rather than glottal stops. More interestingly, four listeners thought each of B7 and B10 were 'probably non-native' (though not the same four listeners for both children), and one listener thought B10 was 'definitely non-native'. Four of these listeners noticed retroflex sounds in the brothers' productions and thought they might be Asian. Retroflex consonants were indeed noticed in the production of B7 and B10 by myself and will be discussed in the relevant chapters. What is interesting though, is that retroflex consonants are not part of Lebanese Arabic and therefore the children must have acquired them from influence outside the home. Note that B10's teacher did mention a Pakistani friend whom he thought had had a great influence on B10's speech.

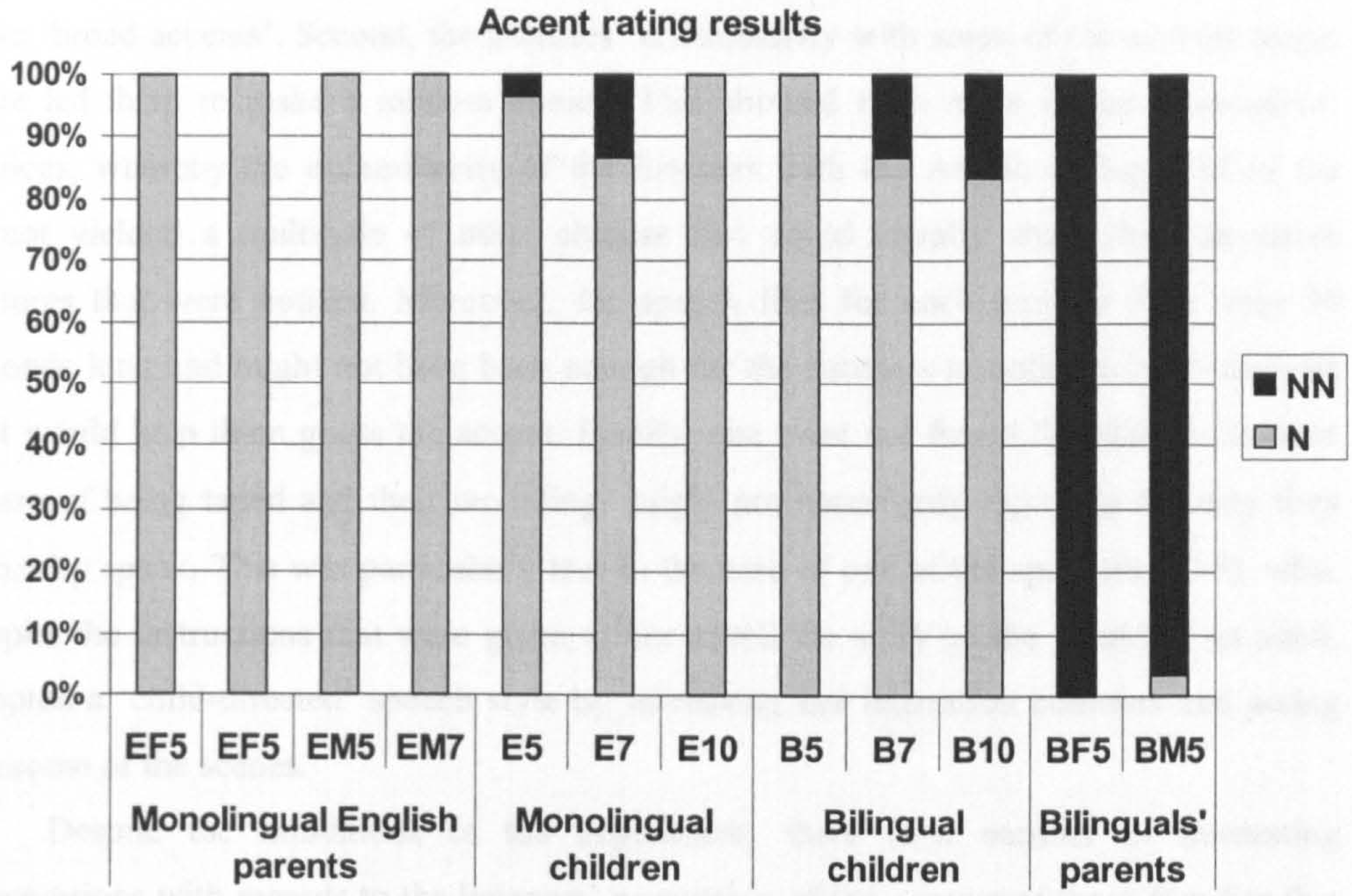


Figure 2.1: Accent rating results for 12 speakers from the study. EF5 was included twice. N = native. NN= non-native.

Table 2.3: Detailed results of the accent rating experiment. N (listeners) = 30.

Language groups		Definitely native		Probably native		Probably non-native		Definitely non-native	
		N	%	N	%	N	%	N	%
Monolingual English parents	EF5	28	**93**	2	**7**	0	**0**	0	**0**
	EF5	28	**93**	2	**7**	0	**0**	0	**0**
	EM5	29	**97**	1	**3**	0	**0**	0	**0**
	EM7	29	**97**	1	**3**	0	**0**	0	**0**
Monolingual English children	E5	12	**40**	17	**57**	1	**3**	0	**0**
	E7	15	**50**	11	**37**	4	**13**	0	**0**
	E10	26	**87**	4	**13**	0	**0**	0	**0**
Bilingual children	B5	24	**80**	6	**20**	0	**0**	0	**0**
	B7	15	**50**	11	**37**	4	**13**	0	**0**
	B10	15	**50**	10	**33**	4	**13**	1	**3**
Bilinguals' parents	BF5	0	**0**	0	**0**	7	**23**	23	**77**
	BM5	0	**0**	1	**3**	8	**27**	21	**70**

With respect to naming the accent of the speakers, the listeners' answers were quite varied and, in many cases, reflected the varied background of the subjects. However, there were also signs of random guesses by the listeners, which may be due to many reasons. First, some of the monolinguals' parents observed that their accents were 'mixed' due to having lived in many places as they were growing up, and others underlined the fact that they came from middle class backgrounds and therefore did not have 'broad accents'. Second, the listeners' unfamiliarity with some of the accents might have led them to make a random choice. This showed even more in the 'non-native' choices, whereby the unfamiliarity of the listeners with the Arabic background of the parent yielded a multitude of other choices that could equally share the non-native features that were noticed. Moreover, the speech files for each speaker were only 30 seconds long and might not have been enough for the listeners to notice enough features that would help them guess the accent. Finally, one must not forget that the adults were aware of being taped and their recordings might not necessarily represent the way they normally speak. This was particularly true in the case of one of the speakers (EF5), who, despite the instructions that were given to her to tell the story as she would to an adult, adopted a 'child-directed' speech style by increasing her intonation contours and acting out some of the scenes.

Despite the limitations of the experiment, there is a number of interesting observations with regards to the listeners' perception of the accents of these families that have been living in Yorkshire for at least 10 years. These will be discussed below, while the detailed answers to the accent-naming experiment and the features observed have been compiled in Appendix five.

First, a lot of the choices for the monolingual parents' accents by the listeners reflect the parents' own perceptions of their accent (Section 2.2). For instance, although EF5 was originally brought up in Northumberland, she observed that her accent generally sounded southern. The majority of the listeners did indeed think her accent was southern, some of them specifying it as 'southern RP', and describing it as 'very well pronounced' (recall that EF5 is the only speaker who told the story as if the audience were children). Interestingly, though, more listeners chose 'northern' when they were played a file for EF5 the second time (EF5 was included twice), but observed that her accent was 'middle class'. EM5, who described his accent as a Leicestershire accent, was indeed recognized by some listeners as being from Leicester or Derby, while others chose nearby towns, 'Midlands', or the general choice 'northern'. There were also choices in the category 'southern', but these were in the minority. EM7 received the most homogenous answers, which reflects his background, as he is one of two subjects in this study who has lived in Yorkshire all his life. Chosen accents included general ones like 'northern', but also more specific ones like 'Yorkshire', 'Leeds', and 'Sheffield'.

With respect to the bilinguals' parents, as mentioned earlier in this section, the unfamiliarity of the listeners with the Arabic dialect, led them to opt for various other non-native backgrounds, including European and Asian ones. However, the features that were highlighted were well-observed and did indeed occur in the speaker's production. These included features like sound substitutions (e.g. /ð/ > [d] and /z/, /θ/ > [s]), tap and trill realizations of /r/'s, close [o] vowels for /əʊ/, syllable-timed rhythm, non-native intonation, hesitations, and factors not related to pronunciation, such as syntactic errors.

Moving to the children, the choices for both groups were quite varied and might have reflected some random guesses, but more importantly, the varied choices reflected the difficulty that was involved in trying to guess the accent of the children. One pattern that emerged included more 'southern' choices for E5, B5, and B10, and more northern choices for E7, B7 and E10. In fact, E10 is the only child who received a variety of 'northern' choices with no southern ones, whereas the choices were quite mixed for the other children.

In sum, the accent-rating experiment reflects the rich sociolinguistic background of the speakers in many ways, not only with respect to the range of native to non-native acquisition of English that they exhibit, but also with respect to their social situation and geographical mobility. What is of major interest to this study, though, is that the overall impression of the bilingual children's accent is that they are native-like in English.

In Chapters 3 to 6, we move on to a detailed description of the sound features chosen for investigation in this study in order to find out whether the analysis supports or contradicts the results from this section.

CHAPTER THREE

/l/ production

3.0 Introduction

This chapter presents results from auditory and acoustic analysis of /l/ in the production of the English and Arabic monolingual and bilingual subjects. In Section 3.1, /l/ production in English is described, taking into account the variety of English /l/ produced in the bilingual subjects' environment and developmental patterns of /l/ acquisition normally found in children. Section 3.2 offers a similar description for /l/ production in Arabic, and Sections 3.3 and 3.4 present what is known about bilingual acquisition of /l/ and sociolinguistic factors that may affect such acquisition. The aims for this chapter are listed in Section 3.5, followed by a description of the material used for /l/ examination and the type of analysis conducted in Section 3.6. The detailed results for the subjects are then presented in Sections 3.7 and 3.8, and a summary of the main patterns follows in Section 3.9.

3.1 English /l/

3.1.1 Articulatory description of English /l/

English /l/ is most commonly a voiced alveolar lateral approximant with a range of clear and dark allophones that are determined by contextual as well as accentual factors (e.g. Cruttenden, 2001; Davenport & Hannah, 1998; Roach, 1991). Articulatory descriptions of the difference between clear and dark /l/'s however, differ from one source to another. In traditional accounts, the production of the clear [l] involves a front articulation only, whereby the tongue tip or blade is in contact with the alveolar ridge, and there is another contact between the hard palate and one or both of the lateral edges of the tongue as the front of the tongue is raised in the direction of the hard palate, giving a front vowel resonance to the consonant. The production of the dark [ɫ], on the other hand, is described as involving both front and back articulations, whereby the tip of the tongue makes contact against the alveolar ridge, the front of the tongue is depressed, and the back is raised in the direction of the soft palate, giving a back vowel (or velarised) resonance (e.g. Cruttenden, 2001: 202; Davenport & Hannahs, 1998: 32; Jones, 1972: 176; Laver, 1994: 307; O'Connor, 1973: 56/148; Roach, 1991: 59). More recent articulatory and acoustic data suggest that clear and dark /l/'s involve the same articulations or gestures; the primary distinction between the two is therefore in the greater amount of tongue body retraction and lowering of the tongue dorsum as well as their earlier occurrence relative to the apical gesture in the dark variety (Carter, 1999; Sproat & Fujimura, 1993, Stevens, 1998: 543).

There is disagreement in the literature with regards to whether the nature of the clear/dark relationship should be encoded at the level of the phonology of syllable-structure (intrinsic explanation), or whether it can be accounted for by extrinsic phonetic interpretations (natural and dialect-specific). For instance, there is a tendency for clear [l] to occur in various syllable-initial contexts[2] ([lɪp] 'lip'; [ˈseɪ.lə] 'sailor'; [bləʊ] 'blow'), whereas dark [ɫ] tends to occur in various syllable-final and syllabic contexts ([pi:ɫ] 'peel'; [bʌɫb] 'bulb'; [teɪ.bɫ̩] 'table') (Cruttenden, 2001: 201). This is thought to be the case due to the /l/ articulation involving an apical and a dorsal gesture in English (Sproat & Fujimura, 1993). The apical gesture is consonantal in nature by virtue of involving complete stricture; it is therefore hypothesised to be attracted to syllable margins as other consonants are. On the other hand, the dorsal gesture is vocalic in nature by virtue of involving an open type of approximation; it is therefore hypothesised to be attracted to syllable nuclei. For this reason, it is assumed that the apical gesture will precede the dorsal one in syllable-initial position and will follow it in syllable-final position (Sproat & Fujimura, 1993). Similarly, intervocalic /l/'s differ in terms of clearness and darkness depending on the type of linguistic boundary they occur in (e.g. 'velar', 'healing'), and are often described as having an intermediate quality between the light and dark variants (Sproat & Fujimura, 1993).

However, internal factors on their own cannot explain the occurrence of clear and dark allophones of /l/, since the implementation of clear-dark alternations in liquids has proven to be not only structure-dependent but also dialect-specific (Carter, 1999). Carter notes that the early dorsal gesture which marks the dark quality for /l/ is not necessarily associated with syllable-final position. Evidence can be found in cross-dialectal differences; in certain English accents like that of Manchester and North Wales dark [ɫ] may occur in all positions, whereas in others like Tyneside and South Wales clear [l] may occur in all positions (Carter, 1999; Cruttenden, 2001: 204; O'Connor, 1973: 149; Wells, 1982). The quality of an initial dark [ɫ], however, might not be as dark as that of a final one. Similarly, a clear final lateral can still be phonetically darker than a clear initial lateral. There are shades of clearness and darkness of /l/'s (cf. Carter, 1999, for a discussion of intrinsic and extrinsic phonetic interpretations and Tollfree, 1999, for a discussion of a continuous model for clear and dark [ɫ]'s within articulatory phonology).

---

[2] Other positions like word-final before a vowel or a /j/ ('feel it'; 'will you') are not discussed here.

There is little experimental evidence about the quality of /l/ in Yorkshire dialects. Wells (1982: 371), however, comments that Yorkshire has 'dark-ish' /l/'s in all environments. On the other hand, Stoddart, Upton, & Widdowson (1999: 76) describe Sheffield /l/'s as normally clear throughout, with dark [ɫ]'s only occasionally occurring in final position among middle-aged and old females, although this comment is not based on a detailed empirical study.

3.1.2 Phonotactic distribution of /l/

In English, /l/ has a highly restricted and predictable context. In syllable onsets, /l/ must occur adjacent to the nucleus. Where the onset consists of a consonant cluster, /l/ can cluster with six different obstruents: /p/, /b/, /k/, /g/, /f/, and /s/ (Cruttenden, 2001: 201). In syllable codas, /l/ must also occur next to the nucleus. /l/ can also be realised as the nucleus of unstressed syllables, e.g. [bɒtl̩] 'bottle' and [pʰɪkl̩] 'pickle' (Cruttenden, 2001: 360).

Apart from the clear and dark varieties of /l/, there are some dialects of English where syllable-final /l/ is vocalised and is realised as a non-syllabic back vocoid, mainly [ʊ], but also [ɤ], [ö], [ɔ̈], [ʉ], [ʌ̈], and [ɐ] e.g. [mɪʊk] 'milk', [fɪɤ] 'fill'. Ladefoged & Maddieson (1996: 193) note that in such vocalised productions, alveolar contact is completely missing (though it is always possible that the apex may be slightly raised) so that the tongue tip is behind the lower front teeth and the tongue back is raised to produce a segment that is acoustically similar to [ʊ]. However, as can be noticed from the various realisations above, not all the vocalised productions are of the rounded variety. Furthermore, the variation between vocalised and dark [ɫ] is thought to be non-categorical and sensitive to articulatory and perceptual factors. For instance, Hardcastle & Barry (1985: 41) note that the occurrence of /l/ vocalisation is influenced by preceding and following contexts, whereby there is a significant preference for its occurrence when followed by [-FRONT] (velar or palato-alveolar) rather than [+FRONT] consonants (e.g. 'milk' *versus* 'milled'), and when preceded by front rather than back vowels (e.g. 'milk' *versus* 'bulk').

In describing /l/ vocalisation as a developing British innovation that was 'very much in progress', Wells (1982: 258) noted that speakers were being inconsistent in their use of the vocalised form as opposed to the lateral approximant. More recently, Tollfree (1999: 174), has noted that /l/ vocalisation in the accent of South East London English varies depending on speakers and context. Younger speakers tend to produce more vocalisation than older ones, and their use extends to contexts not formerly subject to vocalisation in Wells' 1982 description, such as word-final intervocalic ones (e.g. [ligɤwʷɪnfɐɤ] 'legal info'). Tollfree suggests that the higher incidence of vocalisation among the younger group and its recent extension to other contexts is indicative of change. Such an observation is echoed by Cruttenden (2001: 203-204), who notes that although /l/ vocalisation is mainly a feature of Cockney, it is also spreading to London regional RP, especially when a consonant with a labial articulation precedes ('careful', 'people'), but not when other consonants precede ('uncle' 'special'), or after alveolar plosives ('little', 'middle'), as these are considered childish pronunciations. More recent research shows that /l/ vocalisation is also rapidly spreading to other English accents, especially among the young generation (e.g. Williams & Kerswill, 1999: 148).

There are barely any recent comprehensive accounts of Yorkshire dialects, so there is no reliable information on the use of this feature in the area, but traditional descriptions of the area mention forms of 'old', cold' and 'shoulder' as being produced 'without /l/' (Petyt, 1985: 219), though the Petyt found only a handful in his data with no visible pattern of use and with informants who would be described as 'broad'. Another pattern that Petyt found was a diphthongal form of high vowels preceding final /l/'s (e.g. [fuɪl] 'fool'; [wɪəl] 'wheel'). The author notes that the frequency of their use varies depending on the lexical item involved, with [ɪə] for e.g. being used frequently in words like 'field' and 'wheel' and less frequently in words like 'meal' and 'steel'. No remarks are made on the quality of /l/ following the vowels. Both the absence of /l/ in words like 'old' and diphthongisation are mentioned under Petyt's section on 'non-standard features'.

3.1.3 Acoustic description of English /l/

Voiced lateral approximants are normally characterised by relatively weak formant-like resonances due to a narrower constriction than that normally made for vowels, manifested by an average of 10dB lower amplitude in the lateral compared with a following stressed vowel (Johnson, 1997: 155; Stevens, 1998: 534). The first formant is normally low in frequency (typically below 400Hz (4.01Z)[3] for males) and is also prone to abrupt changes at the release into a stressed vowel due to the rapid change in the cross-sectional area of constriction at the lateral release. The second formant stretches over a wide range of frequency depending on the location of tongue closure (900-1600Hz, i.e. 7.91 to 11.52Z), but is typically well separated from the third formant, which has a relatively strong amplitude and a high frequency in the 3000Hz (15.68Z) region (Cruttenden, 2001: 203;

2 All the bark measures are my own addition in order to allow the results from this study to be comparable with other studies, most of which present formant frequency measures in Hz.

Ladefoged & Maddieson, 1996: 193; Stevens, 1998: 546). Apart from the formants, the lateral is usually characterised by an anti-formant in the output spectrum between F2 and F3 due to the creation of a small pocket of air on top of the tongue, which introduces a side branch to the main acoustic channel around one or both sides of the tongue. This pocket resonates at around 2125Hz (13.42Z), and since it is a side cavity, it becomes an anti-resonance in the output at the lips and causes the amplitude of all the higher formants to be reduced by about 1.6 dB (Johnson, 1997: 155). The basic acoustic distinction between the clear and dark varieties of /l/ lies in F1 and F2 frequencies, whereby clear [l] has a relatively high F2 and a low F1, whereas dark [ɫ] has a lower F2 and a higher F1 (Bladon & Al-Bamerni, 1976: 146). In his study of English liquids as produced by 15 adult male speakers of Southern British English, Nolan (1978: 30) reports mean frequency values of 360Hz (3.63Z) for F1 and 1350Hz (10.40Z) for F2 in initial position. Though the mean value for F3 is reported as 3050Hz (15.79Z), Nolan notes that the 'real' F3 was often weak and that the figures derive from what is theoretically F4.

3.1.4 Acquisition of /l/ by monolingual speakers

Very few studies have focused specifically on the development of liquids, and the information gathered here for the acquisition of /l/ is taken from more general studies of phonological development (Bernhardt & Stemberger, 1998: 306/331-334; Cruttenden, 2001: 204; Edwards, 1973: 9/22; Gibbon, 1999; Ingram, 1979: 136-140; Matthews, 2001: 216-218; Menyuk, 1971: 80-83; Moskowitz, 1970; Sander, 1972: 62; Smith, 1973; 2/15; Vihman, 1996: 219-239).

In English, the production of liquids emerges relatively late, preceded by early production of nasals, plosives, and some of the fricatives. /l/ is not regularly present in children's production until the age period 3;0-3;6, during which it is highly variable, and is not normally mastered until the age of 6. The main patterns that appear in children's production in English are gliding and vocalisation, mainly [w] or [ʊ] for both clear and dark /l/'s e.g. [bwu] 'blue'; [ˈwɪəwi]; 'really' [bɔʊ] 'ball', but also [j] for clear /l/, e.g. [ˈjɪzi] 'Lizzie', [ˈjɛjoʊ] 'yellow' (note consonant harmony effects in 'yellow' and 'really'). Vocalisation is most common in English where syllabic consonants occur, and the most common substitution for dark [ɫ] is a back rounded vowel, either [o] or [u] (e.g. 'apple' [apo]), although open unrounded vowels tend to occur as well (e.g. 'wheel' [wɪɐ]). Vocalisation an also be realised as lengthening of the previous vowel, e.g. 'bell' [bɛː]; 'elbow' [ɛbuː]. Other processes include omission, especially in initial consonant clusters, post-consonantal and word-final position, e.g. 'flowers' ['faʊəz]; 'clown' [kaʊn]. Less common substitutions include nasals, fricatives, and stops, e.g. 'Lee' [ni]; 'laugh' [zaf]; 'leaf' [tif]. Clear [l] normally appears before dark [ɫ], as consonants tend to appear in onsets before codas, but also because dark [ɫ] is more prone to gliding and vocalic substitutions.

3.2 Arabic /l/

3.2.1 Articulatory description of Arabic /l/

Arabic /l/ is normally described as a voiced dental or apico-alveolar lateral (Al-Ani, 1970: 48; Shaheen, 1979: 176). Contact is typically made between the tongue tip and the alveolar ridge or teeth, and the front of the tongue is raised towards the hard palate as for a front close vowel [i], while the back of the tongue is depressed in relation to the roof of the mouth (Anani, 1985: 129). Arabic /l/ is clear in all word positions (e.g. [li:fe] 'sponge' and [fi:l] 'elephant'[4]), apart from when it is found in emphatic environments (discussed below). In fact, articulatory descriptions of the Arabic clear [l] mention an apical gesture only, as opposed to the apical and dorsal gestures sometimes described for English /l/. The clear [l] with a single apical gesture has also been used to describe French, German, Hindi and Spanish /l/'s (Cruttenden, 2001: 204; Laver, 1994: 308), and is sometimes called 'flat' [l] due to the lack of a dorsal gesture (Sproat & Fujimura, 1993: 310). Evidence from x-ray tracings for the German /l/ shows a wide unobstructed pharynx with a single apical gesture and a low tongue position below the palate area (Ladefoged & Maddieson, 1996: 184).

Dark [lˁ] in Arabic occurs in extremely limited environments, mainly (1) in an emphatic context if preceded or followed by a back vowel e.g. [lˁɑbɑtˁ] 'he kicked'; (2) in words involving the name of God e.g. [ɑlˁlˁɑ]; and (3) in unpredictable words, sometimes loan words e.g. ['lˁɑmbˁɑ] 'lamp'. Anani (1985: 130) mentions another environment before or after a uvular plosive (e.g. [ħɑlˁq] 'throat') or a uvular fricative ([bɑɣlˁ] 'mule'), but this environment does not apply to the Lebanese dialect under examination due to the absence of the uvular plosive /q/ and the fact that the uvular fricatives are mainly produced as velar.

---

4 No dental diacritic will be used for Arabic /l/ in this chapter as its place of articulation may vary across speakers, dialects, and contexts, e.g. plain *versus* emphatic (Shaheen, 1979).

It is difficult to offer a uniform articulatory description of dark [lˤ] in Arabic due to the disagreement in the literature on the nature of emphatic sounds in general and the articulatory features that they involve. The majority of researchers who dealt with emphatics in the past believed that emphasis is a secondary articulation that is realised as velarisation (e.g. Anani, 1985: 130; Ferguson, 1956: 446; Gairdner, 1925: 15-20; Nasr, 1966; Obrecht, 1968; Omar, 1973). However, most of these researchers used the term 'velarisation' loosely while describing features of emphatics that were clearly pharyngeal, or expressed beliefs that were not based on much experimental evidence (see Laufer & Baer, 1988 for a comprehensive review). A typical description of a velarised [ɫ] is one in which the tip of the tongue is raised to make contact with the alveolar ridge while the back of the tongue is raised towards the soft palate as for the close rounded vowel /u/ (Anani, 1985: 130). Other contemporary and more recent studies have since provided evidence against velarisation and suggested pharyngealisation or even uvularisation (e.g. Adnan Zawaydeh, 1998; Al-Ani, 1970; Delattre, 1971; Harrell, 1957; Jakobson, 1957; Kahn, 1975; Laufer & Baer, 1988; Lehn, 1963; McCarthy, 1994). Still, even the description of pharyngealisation varies greatly from one source to the other in terms of which secondary articulators and which parts of the pharynx (upper or lower) are involved.

The disagreement over articulatory descriptions of dark [lˤ] is largely due to intradialectal and individual variation, but also to the possibility of co-occurrence of more than one feature in the production of emphatics. Lehn (1963: 30), for example, notes the possibility of cooccurrence of any of the following features for Cairene Arabic:

> "1) slight retraction, lateral spreading, and concavity of the tongue and raising of its back (velarisation) 2) faucal and pharyngeal constriction (pharyngealisation) 3) slight lip protrusion or rounding (labialisation), and 4) increased tension of the entire oral and pharyngeal musculature resulting in the emphatics being noticeably more fortis than the plain segments."

In their description of the same Egyptian variety, Kahn (1975) and Shaheen (1979) concentrate only on the pharyngeal aspect of emphatics. Shaheen (1979: 164), for instance, describes dark [lˤ] as a post-dental pharyngealised lateral, though he later notes that the tongue makes a post-alveolar contact, the front of the tongue is depressed, and the root of the tongue assumes the shape of a bulge and is drawn back toward the vertical back wall of the pharynx to form a stricture. Jakobson (1956) and Watson (1999), on the other hand, emphasise the importance of both pharyngealisation and labialisation in the production of Sanʕani and Yemeni emphatics. Laufer & Baer (1988: 193), who examined speakers from different dialects, confirm the pharyngeal aspect of emphatics but offer further detail about the place of constriction. Their fiberscopic analysis highlights the importance of the epiglottis, which forms a constriction with the pharynx walls while the root of the tongue moves backward at the bottom of the pharynx. The authors therefore claim that it is the lower rather than the upper part of the pharynx that is involved in the secondary articulation and that the primary place of articulation remains the same (as opposed to Shaheen's (1979) suggestion that plain [l] is dental while dark [lˤ] is post-dental). McCarthy (1994: 219) and Adnan Zawaydeh (1998), on the other hand, claim that emphatics in all dialects have a constriction in the upper pharynx and that the pharyngealised consonants in Arabic should be called 'uvularised' due to the importance of both the dorsal and the pharyngeal gestures.

Regardless of the place of articulation adopted for clear and dark /l/'s, it is important to note that, similarly to English, there is a definite gradience involved. Mitchell (1993: 25-27) gives an example of a set of words produced with /l/'s ranging from clear to dark along a continuum of darkness: [na:l] 'he obtained' [ʔa:l] 'he said' [tˤɑ:lˤ] 'he reached' [ʔɑlˤlˤɑh] 'god'. Mitchell (1993: 28) also notes the possibility of clear and dark /l/'s being in free variation depending on the extent of emphasis spread and dialectal differences e.g. [ˈbɑtˤɑlˤ] or [ˈbɑtˤɑl] 'hero'.

Apart from the disagreement on the articulatory description of emphatics, their phonological relationship with their plain counterparts is also debated. One problem stems from the identification of different pairs in different sources. The normally undisputed ones are the following pairs, where the emphatic sounds in each pair are called primary emphatics: /t tˤ/; /d dˤ/; /z zˤ/ (or /ð ðˤ/); and /s sˤ/. Then, depending on which source is consulted, some or all of the following pairs might also be listed, with the emphatic sounds considered as secondary emphatics: /k q/; /g ɢ/; /x χ/ or /χ χˤ/; /ɣ ʁ/ or /ʁ ʁˤ/; and /l lˤ/. The choice of pairs often depends on the inventory of a given dialect and the phonological status of emphatic consonants that remain unpaired (see Lehn, 1963 and Laufer & Baer, 1988 for a discussion). Other researchers have noted that labials like [bˤ] [mˤ] and [nˤ] are also characterised by emphasis in the environment of back vowels e.g. [bˤɑ:bˤɑ] 'father'; [mˤɑ:mˤɑ] 'mother'; [nˤɑ:r] 'fire', and can be contrastive with their plain counterparts e.g. [bˤɑ:bˤɑ] 'father' *versus* [ba:ba] 'her door' (Abu Haidar, 1979).

Another problem is whether to consider emphasis a segmental property of the consonant, the vowel, or a suprasegmental prosodic feature affecting consonantal and vocalic articulations. In fact, the perceptual domain of emphasis is at least CV and not just the consonant. Emphasis often affects the following syllable and may even cross syllable boundaries to affect entire words and adjacent words depending on the dialect in question (e.g. Adnan Zawaydeh, 1998; Davis, 1995; Haddad, 1984; Mitchell, 1993; Watson, 1999). More importantly, the phonetic variation of emphasis has been attested on stylistic and phonological levels (Harrell, 1957; Kahn, 1975; Mitchell, 1993). With respect to stylistic differences, both Harrell and Kahn have shown that there is strong cultural and behavioural evidence for the gradient nature of emphasis, reflected in the way a non-emphatic pronunciation of an emphatic sound is reported as being 'affected and effeminate' while an emphatic form sounds 'formal, pompous, or crude and hick-like' (Kahn, 1975: 41).

With respect to the disagreement on the phonological level, the contrast between the clear and dark counterparts of Arabic /l/ for example is sometimes argued to be phonological due to the existence of minimal pairs like [talle] 'hill' and [tˤɑlˤlˤe] 'appearance' (Abu Haidar, 1979; Ferguson, 1956). However, counter-arguments include the fact that those pairs are small in number and the fact that the contrast is in the plain and emphatic stops like /t tˤ/ and not the liquid; the liquid therefore has perseveratory coarticulatory darkness (e.g. Cantineau, 1960: 51).

Among the studies mentioned above, Haddad (1984), Nasr (1966) and Obrecht (1968) are the only three that describe the Lebanese variety. Haddad (1984) concentrates on the phonological aspect of /l/ by stating that emphasis is a suprasegmental feature that affects consonants and vowels alike, mainly causing backing and sometimes rounding of vowels e.g. [lɪsa:n] 'tongue' versus [lˤusˤsˤɑ:n] 'two thieves'. He further notes that [lˤ] in Lebanese Arabic is not an emphatic phoneme, but one that acquires emphasis in the environment of a dental consonant and a back vowel e.g. [ʔɑlˤlˤɑ:h] 'god' *versus* [bɪllaah] 'in the name of god' (note that most examples he gives are expressions related to the name of God). Nasr (1966) and Obrecht (1968) concentrate on the articlulatory aspect of /l/'s, mainly stating that the dark allophone is velarised [lˠ]. Both studies have their weaknesses: Nasr's description is purely perceptual and mentions no acoustic analysis although he at least attempts to describe of the colloquial Arabic of Lebanon; Obrecht, on the other hand, does includes instrumental analysis but his analysis leads him to conclude that emphasis is produced by a constriction in the pharynx, which contradicts with the term 'velarised' that he used. Moreover, his Lebanese subjects read material in Modern Standard Arabic, which does not necessarily represent the colloquial features of the Lebanese dialect. For this reason, the term 'emphatic' in this study will be used to mean pharyngealised rather than velarised due to the ample evidence from the experimental studies mentioned above. In fact, Laufer & Baer (1988) had one Lebanese subject among the 9 subjects in their study, and the strength of their methodology lies in the use of spectrographic and endoscopic observations of real words and sentences alongside nonsense utterances. The authors compared pharyngeal sounds (/ħ ʕ/) with emphatics /zˤ ðˤ dˤ/, and found that emphatics were produced with a constriction at the epiglottis, similar to that of pharyngeals but generally less tight (Laufer & Baer, 1988: 194). The authors did report differences among the consonants and the subjects, but these differences are not discussed in detail.

In summary, clear [l] in Arabic embodies a back cavity shape of a wide unobstructed pharynx and a gradual narrowing of the mouth cavity towards the region of articulatory constriction, while dark [lˤ] has another place of articulation dividing the back cavity behind the alveolar point of articulation.

3.2.2 Phonotactic distribution

Similarly to English, Arabic /l/ occurs adjacent to the nucleus in both syllable onsets and codas. However, as opposed to the restricted context in which English /l/ can occur, Arabic /l/ can cluster with many more obstruents than English, including /b/, /d/, /dˤ/, /t/, /tˤ/, /k/, /ʔ/, /f/, /s/, /ʃ/, and /x/, as well as clustering with other sonorants like /m/ and /n/. Moreover, due to the rich use of inflectional and derivational affixation in Arabic, /l/ can occur as the nucleus of initial (e.g. [l̩ˈbu:me] 'the-owl') and final syllables [ʔalbl̩] 'in-the'. Arabic /l/ is also subject to gemination e.g. [ˈbal:aʃ] 'he started'.

3.2.3 Acoustic description of Arabic /l/

Acoustic analysis of the Arabic lateral reveals significant differences between the spectral characteristics of the pharyngealised [lˤ] and non-pharyngealised [l]. In his investigation of the spectral and temporal characteristics of Egyptian /l/ in all word positions in the environment of long vowels, Shaheen (1979: 167-179) argues that the second formant cavity for [l] is the same as for [i]; both represent a half-wavelength standing wave of the combined mouth-pharynx system behind the articulatory closure. The second formant for [lˤ], on the other hand is dependent on cavities behind and in front of the pharyngeal constriction similar to that of the dependency of F2 on the cavities for [u]. In acoustic terms, one would therefore expect the first formant of [lˤ] to be higher than that of [l] and the second formant to be lower. The author's results confirm his predictions and reveal some other interesting features that distinguish plain [l]'s and emphatic [lˤ]'s. While the spectrum of [l] shows the presence of its three formants at average frequency positions of about 330 Hz (3.33 Z) for F1, 1520 Hz (11.18 Z) for F2, and 2300 Hz (13.94 Z) for F3, the spectrum of [lˤ] is characterised by the absence of clear F3 in all positions investigated. Moreover, [lˤ] has an average frequency of 425 Hz (425 Z) for F1, which is considerably higher than that of [l], and an average frequency of 1045 Hz (8.79 Z) for F2, which is considerably lower than that of [l] (Shaheen, 1979: 179). With respect to the plain [l], the author found that while F1 is slightly higher in initial than in final position, there is little variation in F2 in initial and final position. Such results show that Arabic [l] does not vary depending on word-initial or final position, which suggests that in Arabic clearness/darkness of /l/ does not correlate with syllable or word position. Similarly to [lˤ], F3 for [l] is also absent in the majority of cases, mainly in final position. For final [lˤ], while F1 behaves similarly to that of [l], F2 is lower in final position than in initial one and F3 is missing in all word positions. Below are main frequencies obtained by Shaheen for the three formants for [l] and [lˤ].

Table 3.1: Average steady-state frequency positions in Hz for the first three formants of Egyptian [l] and [lˤ] in initial and final positions adapted from Shaheen (1979: 172-176). Equivalent bark measurements were added in brackets by the present author.

	F1		F2		F3	
	[l]	[lˤ]	[l]	[lˤ]	[l]	[lˤ]
Initial	315 (3.18)	400 (4.01)	1500 (11.09)	1100 (9.11)	2300 (13.94)	
Final	265 (2.66)	295 (2.98)	1500 (11.09)	1000 (8.53)		

As can be seen from Table 3.1, F2 in Arabic [l] does not vary much depending on word-initial or final position, which suggests that in Arabic clearness/darkness of /l/ does not correlate with syllable or word position. Though it is not obvious for all the values in Table 3.1, Shaheen maintains that for both [l] and [lˤ] the frequency of the first two formants are higher in initial than in final position. Shaheen's results with respect to lowering of F2 and raising of F1 for emphatic [lˤ] are echoed in most studies that have conducted acoustic analysis of emphatics (e.g. Laufer & Baer, 1988: 195; Younes, 1993: 135), which the authors attribute to the constriction in the pharynx that accompanies emphatics in comparison with non-emphatic counterparts.

From the descriptions presented in this section, it is clear that the acoustic properties of pharyngealised and non-pharyngealised /l/ are similar to those of velarised (dark) and non-velarised (clear) /l/, as described in Section 3.1.3 for English. Following Shaheen's observation that pharyngealised [lˤ] is characterized by a higher F1 than clear [l], a carefully controlled experiment would in principle allow a comparison between pharyngealised and velarised /l/'s by acoustic means, and therefore potentially distinguish between English and Arabic 'dark' /l/'s. The data from this study, however, did not allow for such an investigation.

3.2.4 Acquisition of /l/ by monolingual speakers

In Arabic /l/, production normally emerges earlier than in English (around the age of 2;0-2;6), reaches an acceptable performance around the age of 3;6, and is mastered around the age of 6 (Amayreh & Dyson, 1998; 2000; Dyson & Amayreh; 2000; Omar, 1973). Dyson & Amayreh (2000: 98) note that the early accuracy of clear [l] in Jordanian Arabic may be due to its high frequency in the language and its relatively high functional load. Developmental processes include /l/ deletion e.g. e.g. [ke:b̥] 'dogs' for adult [kle:b], assimilation, e.g. [ˈħɪwwe] 'pretty' for adult [ˈħɪlwe], and gliding, which is less frequent in Arabic than in English and tends to be restricted to [j] (e.g. [ˈʔajam] 'pen' for adult [ˈʔalam]; [ħaji:b] 'milk' for adult [ħali:b]). Another rare substitution for /l/ in Arabic is [n], although the occurrences are very low and sporadic with respect to age groups and individuals (Dyson & Amayreh, 2000: 109). More importantly and as opposed to most varieties of English, there are no reported cases of /l/ vocalisation in Arabic. Similar observations are normally made about German and French /l/'s, where no labio-velar substitutions take place, and gliding is restricted to [ɥ] in French and [j] in German. Therefore, it seems that the different realisations of /l/ produced by children are motivated by the phonology of the language that they are developing.

As for emphatic [lˤ], the difficulty in its production is related to the general difficulty experienced by Arab children in acquiring emphatics due to the articulatory complexity of these sounds that involve simultaneous articulatory postures (Dyson & Amayreh, 2000: 84; Omar, 1973: 55). The usual pattern that appears in the production of emphatics by children is de-emphasis, i.e. the loss of the secondary articulation and therefore producing the plain counterpart of the emphatic sound in question, e.g. [ta:be] 'ball' for adult [tˤɑbe]; [lati:f] for adult [lˤɑtˤi:f]. Since emphatic [lˤ] in Arabic is mainly produced as a result of an emphatic context (e.g. [bɑtˤɑlˤ] 'hero'), its correct production usually depends on whether or not the other emphatic sound(s) within the same utterance have been acquired. Though the incidence of de-emphasis gradually declines with age, it does not easily disappear and sometimes persists even after the age of six due to the infrequency and low functional load of emphatics in Arabic (Dyson & Amayreh, 2000: 100).

3.3 Bilingual acquisition of /l/

Some of the few studies that have looked at early acquisition of /l/ by bilinguals include more general studies of phonological development like Leopold (1970), Burling (1973), Holm & Dodd (1999), and studies that looked at /l/'s in particular, like Ball, Müller & Munro (2001a), Martinez-Dauden & Llisteri (1990) and Pieras Guasp (2001). Each of these studies will be discussed briefly in this section.

Leopold's (1970) study of his English-German bilingual daughter's production in the first two years of life shows early signs of separation in her /l/ production of each language. Leopold (1970: 116) noted that his daughter was treating English and German /l/'s differently due to the difference in the manner of /l/ production in the two languages. Hildegard articulated the German /l/ with a 'flatter tongue' than English, which Leopold (1970: 64) described as being accompanied with more or less raising of the back of the tongue. While Hildegard treated German and English similarly in initial position and substituted them with [h] or [j], she showed signs of different substitutions for /l/ in final position. For instance, 'ball' was often produced [baɪ] in a German context and [baʊ] in an English context. Leopold interpreted the first production as conforming to the 'flat' (i.e clear) nature of German /l/ and the second one as conforming to the bunched dorsal (i.e. dark) nature of the English /l/. The latter was also considered the reason behind more omissions of English than German /l/'s in Hildegard's production as an attempt to avoid the difficult dorsal gesture. Hildegard's vocalisations in English included not only [ʊ], but also [a], [ə], and [ɐ]. Similarly, her intervocalic /l/'s were often omitted in English, whereas the German ones were never completely omitted, but were produced as plain or velarised /l/'s, and were frequently substituted by [j] (Leopold: 1970: 67). The appearance of a velar quality in German /l/'s was noted in early stages where Hildegard had similar productions for a word with English and German equivalents e.g. 'all' and 'alle' both produced with a dark [ɫ], but the child soon changed the German /l/ pronunciation to a palatalised one. The author also noticed less /l/ velarisation in his daughter's English, but attributed that to German influence along with the articulatory difficulty associated with raising the back of the tongue.

Another case of early differentiation between the patterns of liquid production by a bilingual child is Burling's (1971) study of his English-Garo speaking child between the ages of 1;4 and 2;8. Although the author notes that his son's awareness of the two languages being different only emerged at the age of 2;2, the description of earlier productions of /l/'s and /r/'s in the two languages provides evidence for differentiation. For instance, between the ages of 1;5 and 2;8, Stephen used [l] for both Garo [l] and [ɾ] e.g. [lama] for /ɾama/ 'road' (in Garo the two sounds are allophones of the same phoneme, with [ɾ] occurring in syllable-initial position and a lateral similar to English [l] occurring elsewhere). In English, however, Stephen replaced English /r/ with [w] or omitted it altogether. None of the Garo liquids were replaced with the labial-velar approximant.

Holm & Dodd (1999) found differences in behaviour in the two successive Cantonese-English bilinguals that they examined during a longitudinal study between the ages of 2;3 to 3;1 and 2;9 to 3;5. Both children showed different error patterns for each language and clear signs of phonological differentiation, and their behaviour with respect to /l/ substitutions were different. In Cantonese [l] and [n] act as allophones and are common substitutions in the speech development of monolingual children. In Holm & Dodd's study, one of the two bilinguals, Catherine, substituted [n] for /l/ in Cantonese but not in English, which supports the author's claim that she was using different substitutions in the two languages. The second child Max, on the other hand, did not substitute [n] for /l/ in Cantonese, but did so in English, where such substitution is normally less common in monolingual development. The authors concluded that bilingual children not only acquire their phonologies in ways that are different from monolingual children acquiring each language in isolation, but also differ amongst themselves in their acquisition patterns, in the way their two languages interact and in the way they build hypotheses while trying to select appropriate language-specific realisation rules (Holm & Dodd, 1999: 375).

Ball et al (2001a) examined the developmental patterns in the acquisition of the Welsh lateral fricative [ɬ] in 85 Welsh-English bilingual children between the ages of 2;6 and 5;0 divided into five age ranges and into Welsh-dominant or English-dominant subjects. Although their study is not on the acquisition of the lateral approximant [l], it is reviewed here because their results highlight the importance of taking language dominance into account. Since Welsh is spoken by about half a million speakers in Wales, the authors managed to examine subjects who were mastering both languages simultaneously in a predominantly bilingual environment. The Welsh lateral fricative showed the greatest variation among all the sounds investigated by the authors, with around 20 different variants found in the data, and the amount of variability differed between the two dominance groups and decreased over time. The variants included the lateral fricative [ɬ], a wide range of fricatives with the voiceless velar fricative [x], fricative plus lateral clusters (e.g. [xl], lateral plus fricative clusters (e.g. [lʃ]), among others. The English-dominant subjects had low percentages of correct realisation of the lateral fricative, with the highest proportions in initial position (reaching 81% in the oldest group), and very low proportions in word-medial and especially word-final position (reaching only 50% in the oldest group). The most common category of substitutions was in the fricative range, mainly the velar fricative, and the fricative plus lateral clusters increase with age. Other substitutions included stops, plain [l] (mainly in final position), and clusters like [kl]. As for the Welsh-dominant bilinguals, the percentage of correct realisation of the lateral fricative was much higher than that of the English-dominant group, even among the youngest subjects, and reaching 100% accuracy in the 4;6-5;0 group. Like for the English-dominant group, higher proportions of correct realisations were found in word-initial than in word-medial and final positions. The dominant substitution patterns involved the use of fricatives, mainly [x] and [s], and there was a noted lack of the fricative plus lateral category in initial (though noted in the English-dominant group) and final position, as well as a lack of the 'other' category noted in the English-dominant group. Nearly all substitutions were fricatives, increasing to velar fricatives with age. Ball et al's study showed that differences in rate of acquisition and amount of variability are clearly linked to the dominant language of the subjects, and the use of substitutions derives from acoustic as well as articulatory similarity with the target sound.

Studies on older subjects include Martinez-Dauden & Llisteri (1990), who examined the production of /l/ in French as a third language by seven male Spanish-Catalan bilinguals. The study was conducted due to the fact that French and Spanish /l/'s are clear, while Catalan /l/ is described as having a velarised nature characterised by a lingual retraction similar to the so-called varieties of dark [ɫ]. The subjects were university students who were living in Catalonia and who ranged from mainly Spanish speakers to mainly Catalan speakers. F2 measurement were taken as a correlate for darkness and were made for /l/'s produced by the subjects in vowel-lateral-/e/ environments. Results showed that the subjects as a group produced mean F2 measurements (1579Hz/11.43Z) which were similar to the native Spanish average (1534Hz/11.24Z) and not far off from the French one (1656Hz/11.75Z), but significantly different from the low Catalan average F2 (1039Hz/8.76Z) (note that the measurement for the three languages are not from the same study but are reported by the authors for subjects from other studies for comparative purposes). Since there was no significant difference between the group average F2 production in French and the native average in Spanish, Martinez-Dauden & Llisteri interpreted the results in terms of phonetic interference whereby the subjects were transferring the acoustic feature of the Spanish lateral to their spoken French.

Individual results, on the other hand, were more revealing. On the one hand, they showed that the two subjects who were the most balanced bilinguals were the only ones to show signs of significant difference between their French and Spanish /l/ production. On the other hand, the two dominant Catalan speakers did not show an influence of transfer from Catalan to their French laterals. The authors interpreted the results as due to the Spanish school environment in which the speakers learned their French and the fact that one of these two subjects is the most fluent in French. Finally, the authors compared their results to three other studies that have studied the realisation of Spanish laterals by Spanish-Catalan speakers and that have found interference due to the transfer of the 'velar' character of Catalan into Spanish (Comas, 1986; Huerto, Sabio, Silvestre & Sonia, 1988; Martinez, 1989). They note that the subjects in those studies were not balanced bilinguals. Martinez-Dauden & Llisteri (1990) conclude that 'bilingual' is a designation that embodies different degrees of speech production control, at least when referring to the phonetic abilities of individuals in a community where two languages are spoken.

Another study of /l/ involving Spanish-Catalan bilinguals is by Pieras Guasp (2001), who examined the production of 31 bilingual subjects of different ages from Palma, Spain, where Castilian Spanish and Catalan have been in contact for a long time. The velarised Catalan /l/ is acquiring a stigmatised value and seems to be undergoing change towards the clear variety normally found in Spanish. Although the subjects had Spanish or Catalan as their L1, they all seemed sensitive towards the linguistic change that is taking place in apparent time in Catalan. While 13 of the 14 Spanish L1 speakers produced a light [l] in Spanish, only seven out of the 17 Catalan L1 speakers produced a dark [ɫ] in Spanish, therefore showing signs of interference from their L1. Further analysis showed that those seven speakers were from the older generation of Catalan speakers who had learned Spanish as a second language before the 1960s, and therefore before Spanish dominated formal spheres of language use for political and historical reasons (see Pieras Guasp, 2001: 164-165 for discussion). Those speakers had less socialisation with Spanish speakers and the author describes their phonological system as having been 'fossilised' before the spread of the light [l]. The 10 Catalan L1 speakers who produced the correct light variant [l] for Spanish were from a younger generation and were exposed to more frequent social contact with Spanish monolinguals. More interestingly, eight out of these ten speakers also produced a light [l] in Catalan too. Pieras Guasp (2001: 166) explains this phenomenon in terms of the prestige form, which is the light [l], taking over the dark variety used in traditional Catalan. More importantly, the young bilinguals seem to be aware of the stigma and 'notoriety' associated with velarisation, and are becoming part of the change that is affecting Catalan /l/ and driving speakers towards the adoption of the innovative light [l] pronunciation.

Studies of second language learning in children can also offer insight into how developmental factors interact with the introduction of a new language and are likely to cause transfer/interference from the native language to the second, at least in the initial acquisition stages. Hecht and Mulford (1982) found that their subject, an Icelandic boy aged 6;0 learning English in a natural setting, was facing pronunciation difficulties in the initial stages, mainly due to the differences between English and Icelandic. Even though the subject acquired English rapidly, his pronunciation of the language retained several noticeable phonetic characteristics that were transferred from Icelandic, such as a trilled [r] and a clear [l]. Snow and Hoefnagel-Hohle (1977) also found that children as well as adults do not achieve native-like pronunciation patterns when learning a language in a naturalistic setting, although they might achieve a better standard than adults.

There is barely any research about how English-Arabic bilingual children acquire /l/ patterns in either language, although there are important differences with respect to phonotactic constraints and phonological patterning in each of the languages.

3.4 Sociolinguistic issues in the bilingual acquisition of /l/

Very few studies have considered the phonological repertoire of bilingual children with the particular local accent(s) spoken in their environment in mind in order to examine the motivating factors that trigger the production of one realisation over a number of competing alternatives. With respect to /l/, the descriptions in Sections 3.1 and 3.2 have shown that there are important language- and dialect-specific factors that are involved in its production. These factors should be taken into consideration when deciding whether a bilingual child has acquired the appropriate patterns for each language.

Studies on child second language acquisition in a natural setting can be revealing with respect to the extent of acquisition of accent features of a host community. One such study is by Verma, Firth, & Corrigan (1992), who examined the developing phonological system of Panjabi/Urdu speaking children learning English as a second language in two different dialect areas in Britain, West Yorkshire and Scotland. The differences between phonological features in the children's mother tongue and those of the two British varieties helped the authors tease out features in the children's production that are due to L1 interference from ones that are particular to the accent spoken in the subjects' locality or ones that are simply developmental. For instance the Edinburgh subjects were reported using velarised [ɫ]'s in English due to the influence of their local variety, since /l/ is clear in all positions in Panjabi and Urdu. Note however, that the only example given is 'bill' [bɪɫ], with a final dark [ɫ] typical of many British varieties; a stronger argument could have been made of the subjects were noted producing initial dark [ɫ]'s typical of the Scottish variety (Wells, 1982; Stuart-Smith, 1999). Not much information is available on the behaviour of the Yorkshire subjects with respect to /l/, since the target Yorkshire variant is listed vaguely as the clear [l] variety with no discussion of contextual allophones and their realisation in Yorkshire or by the subjects. Note that initial dark [ɫ] is possible for the Yorkshire variety (e.g. Wells, 1982).

Heselwood & McChrystal (2000) investigated the presence of Panjabi accent features in the English of 19 ten-year-old bilingual children as perceived by 45 phonetically-trained listeners. Among the features examined were clear allophones of /l/ in syllable codas, front epenthetic vowels as the nucleus of otherwise syllabic liquids and nasals in English (e.g. 'candles', 'garden'), and postvocalic /r/. Although more males than females were rated as having non-English accent features in general, both males and females who were given high accent scores had a high incidence of clear allophones of /l/ that were described as 'much clearer' than the norm for realisations of /l/ in coda positions, and a similar number of epenthetic front vowels of a quality that is different from the schwa-type vowel found in monolingual English in syllabic environments (Heselwood & McChrystal, 2000: 51). The authors concluded that the features that the bilinguals exhibited can most easily be explained as influences from Panjabi, the L1 of the subjects, since clear [l]'s in all positions and the absence of syllabic laterals and nasals are features of Panjabi. The subjects who produced them also had a high incidence of retroflex articulations for stops (the males more so than the females), which is a feature that is most noted by the listeners as inducing a strong accent and is associated with Panjabi (Heselwood & McChrystal, 2000: 57).

The two studies reviewed in this section underline to importance of determining the targets that are available to the bilingual from their local community before judging their productions. Although clear and dark varieties of /l/ do not contribute towards lexical contrasts, they constitute an important part of the sociolinguistic acquisition by the speakers.

3.5 Aims of the study

In light of the preceding discussion which has drawn attention to the importance of taking social dimensions into consideration when defining a 'phonological system' and of the role of the language mode in analysing bilingual data, this chapter examines the extent to which bilingual children can establish phonetically/ phonologically distinct production patterns for /l/ in each language. The experiment is designed to investigate the following questions:

1 Do English-Arabic bilinguals acquire separate /l/ production patterns for each of their languages?
2 Are their patterns of production in each language similar to those of the monolingual controls in the study?
3 Are the patterns for the monolingual subjects in this study similar to the ones normally described in the literature and therefore expected for each language?
4 Are there signs of influence from one language onto the other in the bilinguals' production and what are the factors that affect such influence?

3.6 Procedure

3.6.1 Material collected for the /l/ variable

Data for this chapter are taken both from recordings from the Leeds IViE (Intonational Variation in English) corpus (Grabe & Nolan, 2001) and the recordings collected for this study. As mentioned in Section 3.1, there are hardly any up-to-date accounts of the quality of /l/ found in Yorkshire dialects, so the availability of Grabe & Nolan's recordings of speakers from Leeds was very helpful in that it offered data on present-day Leeds speech, even if from a small sample of speakers. Since the data were collected for the analysis of a different linguistic purpose (for the study of intonation), the only way to collect enough /l/ tokens for each of the speakers in word-initial and word-final position was to combine data from three different speech styles in which the speakers were recorded (reading passage, free conversation, and story telling). Around 50 to 55 /l/ tokens per speaker in a variety of vocalic contexts were then available for analysis.

Material from this study was collected from: (i) words produced in isolation during the picture-naming activities for the children and the reading lists for the adults, (ii) running speech during the story telling activities for both children and adults, (iii) free play sessions for the children and (iv) interviews with the adults. With respect to the words produced in isolation, the two contexts chosen for examination were absolute word-initial and word-final positions in order to control for the surrounding contexts, as these are the tokens that were also acoustically analysed. As for the rest of the data, /l/ tokens were chosen from a variety of onset and coda contexts, including clusters and intervocalic positions, but excluding contexts where /l/ might be ambisyllabic (e.g. 'calling') or where re-syllabification might occur (e.g. 'feel it'). In English, word-final tokens were further divided into word-final and syllabic in order to examine any difference in the amount of /l/ vocalisations between the two contexts. In Arabic, all the /l/ tokens that were produced in an emphatic environment were also extracted from the data (Table 3.2).

Table 3.2: Sample tokens used for the examination of /l/ in English and Arabic

English	**Onset**	**Coda**	
		Non-syllabic	**Syllabic**
Examples	lap sleep fly happily	pool elbow older bell	bottle purple kettle ankle

Arabic	**Onset**		**Coda**		**Emphatic context**	
Examples	**IPA**	**Gloss**	**IPA**	**Gloss**	**IPA**	**Gloss**
	ˈlaban	yogurt	ˈʒamal	camel	ˈbɑsˤɑlˤ	onion
	l̩ˈbuːme	the owl	ˈħɪlwe	elephant	ʕɑdˤɑlˤ	muscle
	ˈtleːte	three	ˈkalbo	his dog	ˈtˤɑlˤːe	view
	kleb	dogs	ʔajˈluːl	September	tˤɑwlˤɑ	table

3.6.2 Analysis

While only auditory analysis was conducted on the IViE data and on the near-naturalistic data collected for the current study, both auditory and acoustic investigations were conducted on the tokens produced in isolation by the children and the adults from this study. During the auditory analysis, the /l/ tokens that were produced were labeled along a 4-point scale including 'clear', 'medium', 'dark', and 'vocalised'. The medium category was chosen to test reports of Yorkshire /l/'s being of an intermediate darkness quality (e.g. Wells, 1982) and to avoid having to make a forced 'clear' *versus* 'dark' choice for tokens from either language. As for the acoustic analysis, measurements of the first three formants were taken for all tokens, with F1 and F2 frequency being used as the main correlates of clearness/darkness in /l/'s. Formant measurements were made at a relatively steady state in the formant trajectory or at the mid-point of the liquid where there was no evidence of a steady state. Measurements were taken using spectra with 25ms Hanning windows and were double-checked by visual inspection of wideband spectrograms. All formant frequencies were then bark-scaled in order to obtain a perceptual basis for clearness/darkness, since the relationship between perceived quality differences and formant frequency intervals is not linear (Bladon & Al-Bamerni, 1976: 143).

A total of 3161 tokens were analysed for this chapter, consisting of around 446 /l/ tokens from the IViE corpus and 1870 tokens from this study which were auditorily analysed, and another 845 /l/ tokens which were analysed auditorily and acoustically.

3.6.3 Presentation of results

As explained in Chapter Two, initials and numbers will be used for the subjects in the presentation of results (Table 3.3). Numbers will be added to the initials of the adults to help identify their child/children, e.g. EM5; BF7; AM10, etc.

Table 3.3: Initials used for the subjects in the presentation of results.

	Age 5	**Age 7**	**Age 10**	**Adults**	
Monolingual English	E5	E7	E10	EF	EM
Bilingual	B5	B7	B10	BF	BM
Monolingual Arabic	A5	A7	A10	AF	AM
Total = 23	**3**	**3**	**3**	7	7

3.7 English results

3.7.1 Onset position: adults

3.7.1.1 auditory analysis

Results for the adults from the IViE corpus are discussed first, followed by the results from the adults in this study in order to assess in more detail the specific targets that are available for the children.

Figure 3.1 shows the patterns found for the 10 females and males from Leeds the IViE corpus. An important observation from the patterns found in onset position is that the production of initial dark [ɫ]'s is prevalent in some but not all of the speakers. Knowing that all 10 speakers produced approximately the same words (Appendix two), the differences in their production seem to be strongly related to gender. There is a clear tendency for males to use more dark initial [ɫ]'s than females (X^2 test, $p << 0.001$). The amount of dark [ɫ] production stretches from as little as 5 out of 25 tokens for F5 to almost categorical use by M1 (22 out of 23).

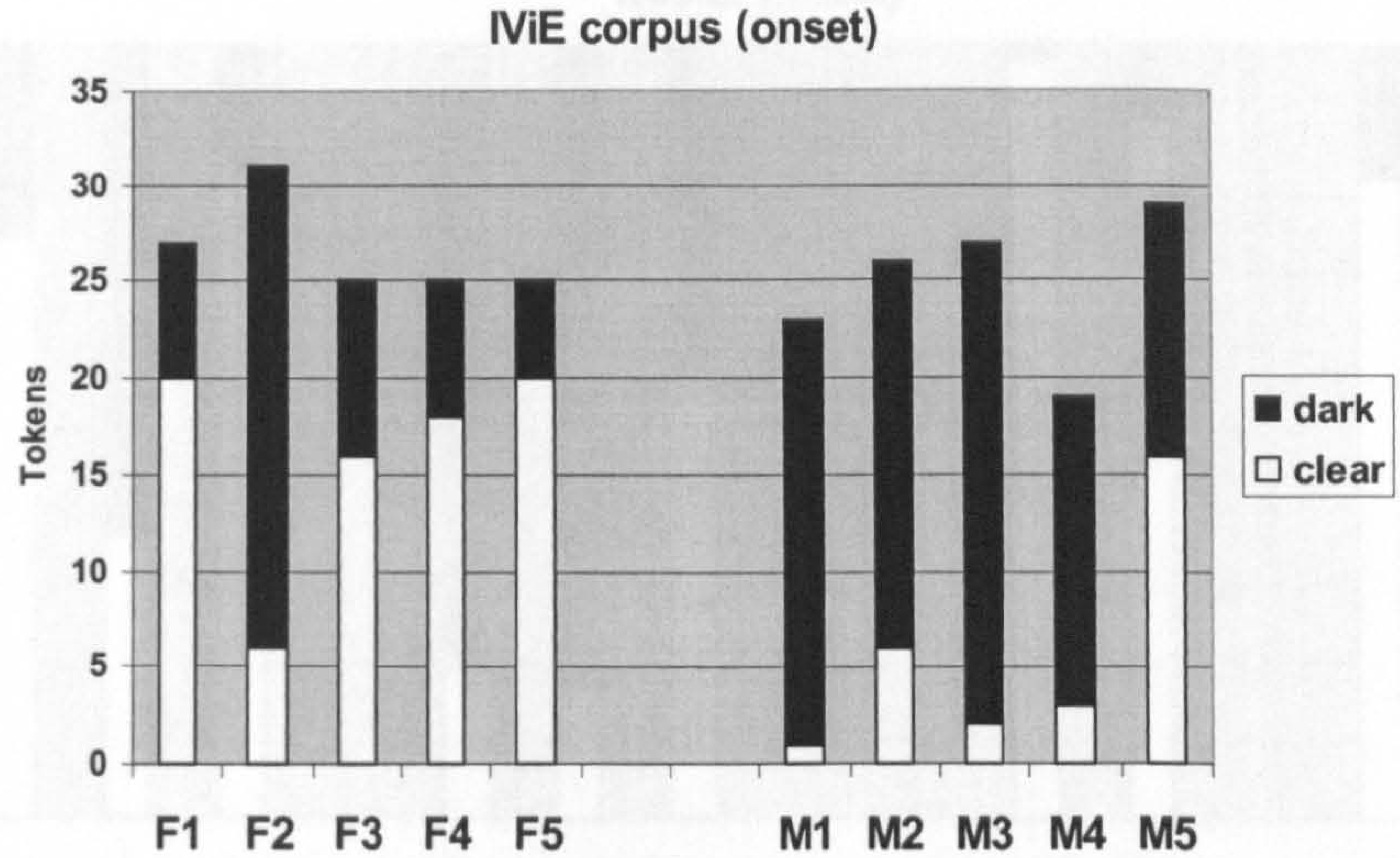


Figure 3.1: Auditory results for syllable-onset /l/'s produced by the monolingual English adults from the IViE corpus. N = 257.

The behaviour of the monolinguals' parents from this study (Figure 3.2 and Table 3.4) is similar to that of the speakers from the IViE corpus. Once again, the differences in the six speakers' production seems to be related to gender, but may also be due to the original accent of the parents. With respect to gender, there is a strong tendency for males to use more dark initial [ɫ]'s than females (X^2 test, $p << 0.001$). The percentage of intermediate and dark [ɫ] production stretches from as little as 16% for EF10 to 83% for EM7. But since EM7 father has lived in Yorkshire all his life, his production might reflect the accent of the area. Therefore, the target for /l/ production that is available for the children in this study is variable and gender-related.

As for the bilinguals' parents, they mainly produce clear [l]'s in this position, with no significant differences between the four speakers. Such a pattern is expected, since the parents have all learned English as a foreign language in their adulthood. A small number of the productions by the bilinguals' parents did fall under the dark category, but these tokens were mainly produced when followed by a back open vowel (e.g. [ɫɑɪt] 'light'; [ɫɑɫɑ] 'La La'), and are therefore likely to be due to effects of coarticulation (Nolan, 1982). Finally, all ten adults subjects were consistent in the /l/ patterns they produced regardless of the style (reading list *versus* story telling and interview data) (Table 3.4).

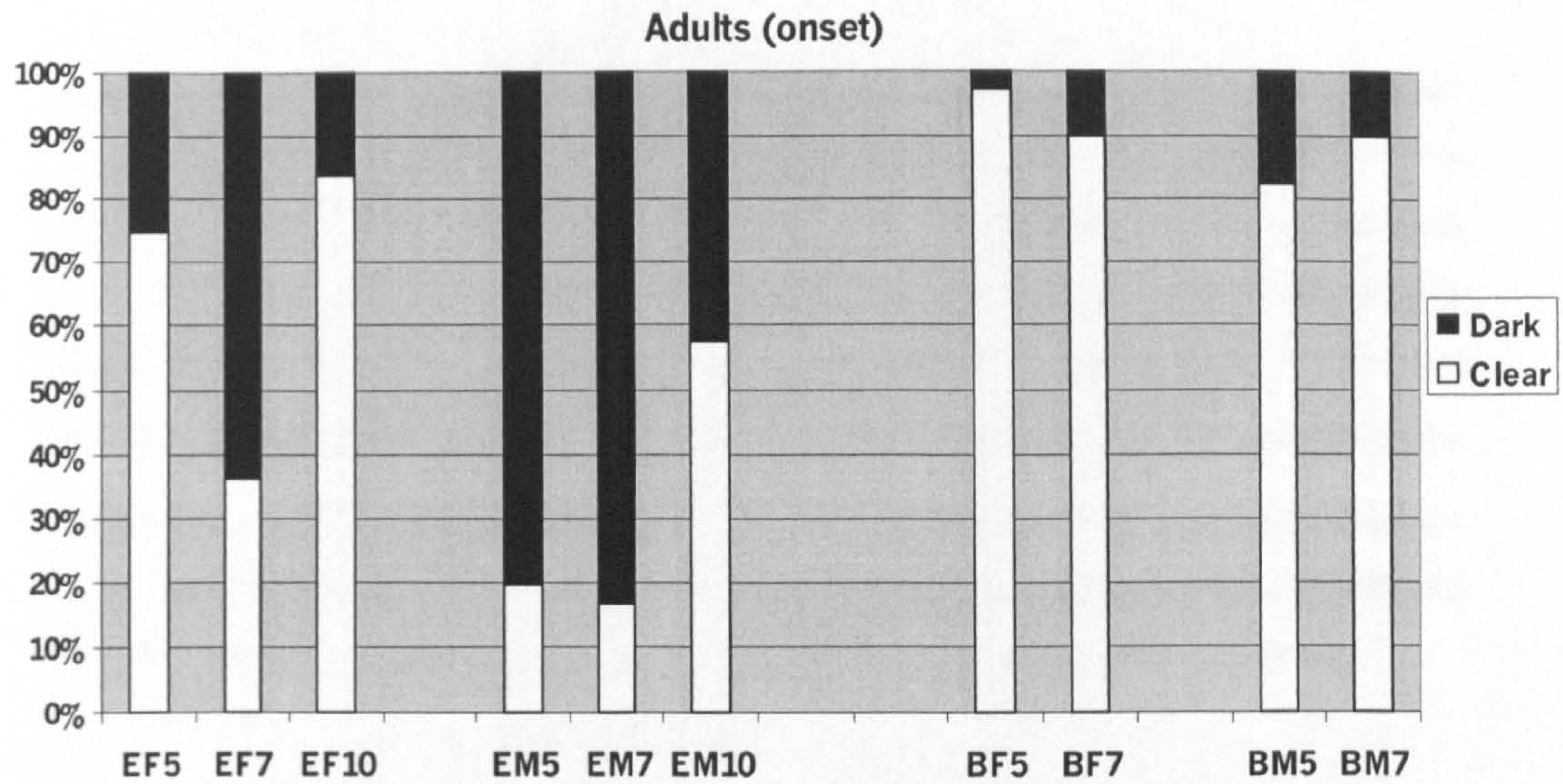


Figure 3.2: Auditory results for /l/'s in syllable-onset position as produced by the monolinguals' parents (left) and the bilinguals' parents (right). Medium and dark categories are combined. N = 685.

Table 3.4: Detailed results for /l/ production in syllable-onset position during the reading list and story telling activities for the adults in English.

	Monolinguals' mothers								
	EF5			EF7			EF10		
	rea d	story	N	read	story	N	read	story	N
Clear	14	39	53	6	38	44	9	69	78
Med	0	10	10	0	22	22	0	6	6
Dark	2	6	8	10	46	56	6	3	9
Total	**16**	**55**	**71**	**16**	**106**	**122**	**15**	**78**	**93**

	Monolinguals' fathers								
	EM5			EM7			EM10		
	read	story	N	read	story	N	read	story	N
Clea r	1	13	14	0	14	14	6	44	50
Med	0	18	18	0	11	11	0	15	15
Dark	15	24	39	16	43	59	10	12	22
Total	**16**	**55**	**71**	**16**	**68**	**84**	**16**	**71**	**87**

	Bilinguals' mothers						Bilinguals' fathers					
	BF5			BF7			BM5			BM7		
	rea d	story	N	read	story	N	read	story	N	rea d	story	N
Clear	17	18	35	14	13	27	13	29	42	12	24	36
Med	0	0	0	0	0	0	0	2	2	0	1	1
Dark	1	0	1	2	1	3	5	2	7	3	0	3
Total	**18**	**18**	**36**	**16**	**14**	**30**	**18**	**33**	**51**	**15**	**25**	**40**

3.7.1.2 acoustic analysis

Acoustic analysis was conducted on tokens produced in isolation in absolute word-initial and final position. Figure 3.3 shows F2 distribution for initial /l/ produced by the monolinguals' parents and the bilinguals' parents, with an indication of whether the token was heard as clear (white dashes) or dark (black dashes). The bark-scaled figures correspond to measurements made from /l/ tokens followed by the vowels /iː/, /eɪ/, /a/, /aɪ/, /ɔ/, or /u/. Since there were not enough tokens for each vocalic context to be presented separately, an equal number of tokens from each context was included for each of the adults in order to obtain comparable data. The measurements for the mothers and fathers are presented separately due to the higher F2 measures expected for females.

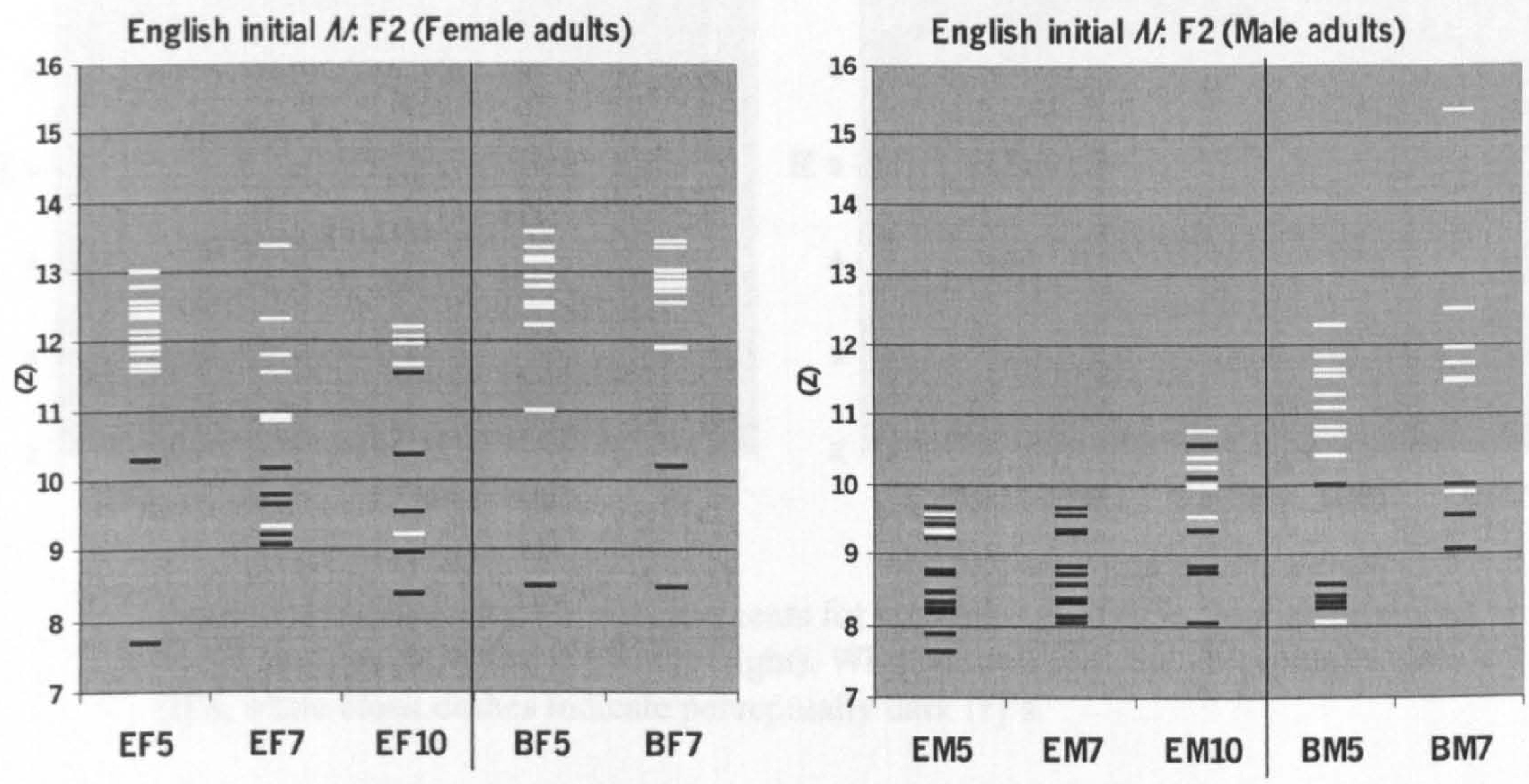


Figure 3.3: Bark-scaled F2 measurements for syllable-initial /l/ in English produced by the monolinguals' mothers (left) and the fathers (right). White dashes indicate perceptually clear [l]'s, while black dashes indicate perceptually dark [ɫ]'s. N = 162.

As can be seen the above figure, dark tokens for all speakers tend to have lower F2 measurements, as predicted by Bladon & Al-Bamerni (1976: 146). EM5 and EM7 produce the lowest and the most concentrated F2 frequencies ranging only between 7.63 and 9.66Z, while the other adults produce a wide range of F2 frequencies. Note that EM5 and EM7 are the speakers who showed the highest use of initial dark [ɫ]'s in Figure 3.1. It is also interesting to note the slight overlap between the measures for tokens labeled clear and the ones labeled dark. While the majority of low F2 measurements belong to tokens that were labeled dark, there are a few clear [l] tokens with equally low F2 (e.g. 8.05Z for BM5, 9.31Z for EF7, and 9.20Z for EF10), and a few tokens with high F2 measurements that were perceptually heard as dark (e.g. 11.64Z for EF10, 10.23Z and 10.27Z for EM10).

As also predicted by Bladon & Al-Bamerni (1976: 146), F1 measurements show a tendency for dark [ɫ]'s to have a higher frequency than clear ones, although the difference is not as straightforward as for the F2 patterns (Figure 3.4). Although the highest F1 measures for all speakers except EF7 belong to dark [ɫ]'s, the overlap between clear and dark /l/'s with respect to F1 frequency is much greater than that for F2.

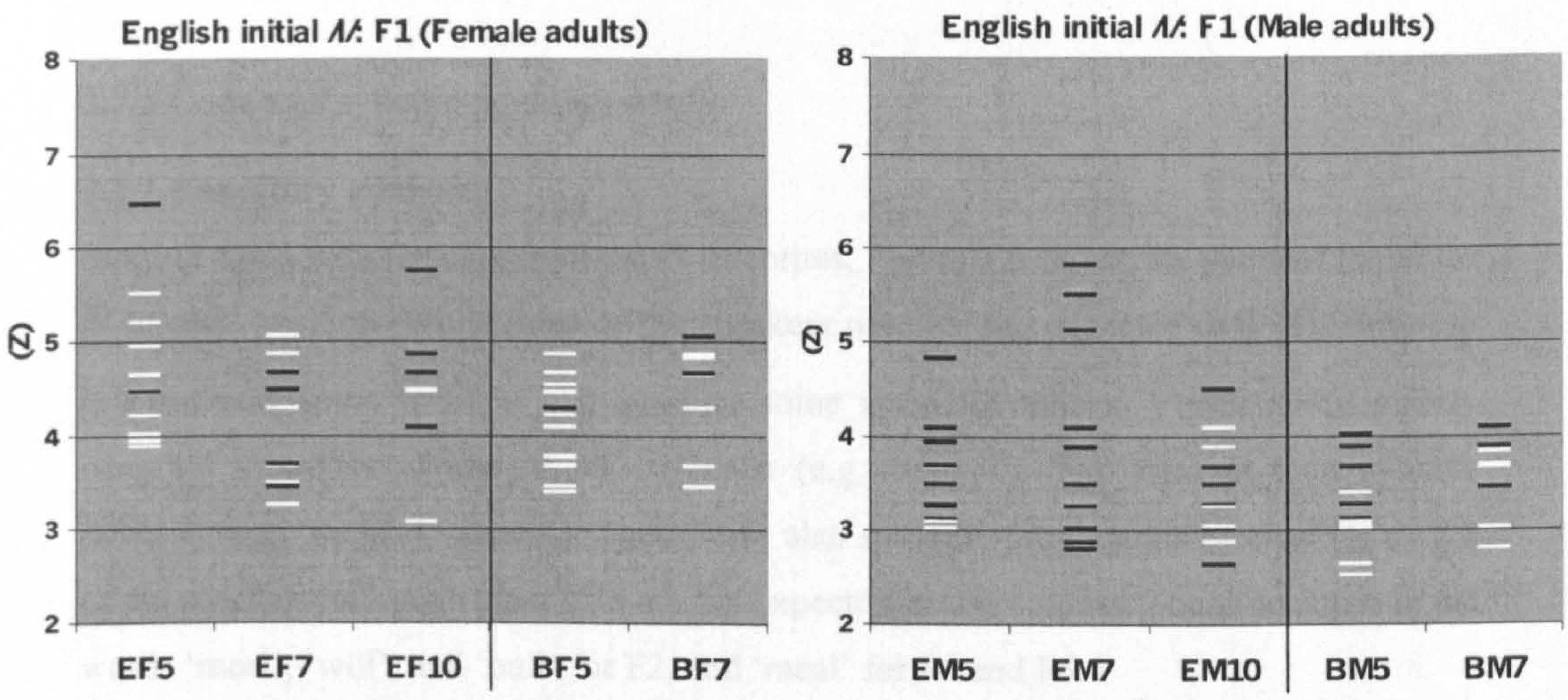


Figure 3.4: Bark-scaled F1 measurements for syllable-initial /l/ in English produced by the females (left) and the males (right). White dashes indicate perceptually clear [l]'s, while black dashes indicate perceptually dark [ɫ]'s.

The patterns just described for /l/ production by the monolinguals' parents and the bilinguals' parents illustrate the type of variability that is available in adult input to the child and provide further evidence for the claim that there often is no stable target model for the child to acquire (e.g. Foulkes et al, 1999; Local, 1983). Figure 3.5 illustrates the kind of acoustic variability that is available in the input when one examines the production of the word 'leaf', which was realised as [ɫif] by one of the monolinguals' fathers (EM7) with a low F2 of 9.54 Z, and as [li:f] by one of the bilinguals' fathers (BM7) with a high F2 of 12.48 Z. In the presentation of the children's results we will examine in what way such variability shapes their productions.

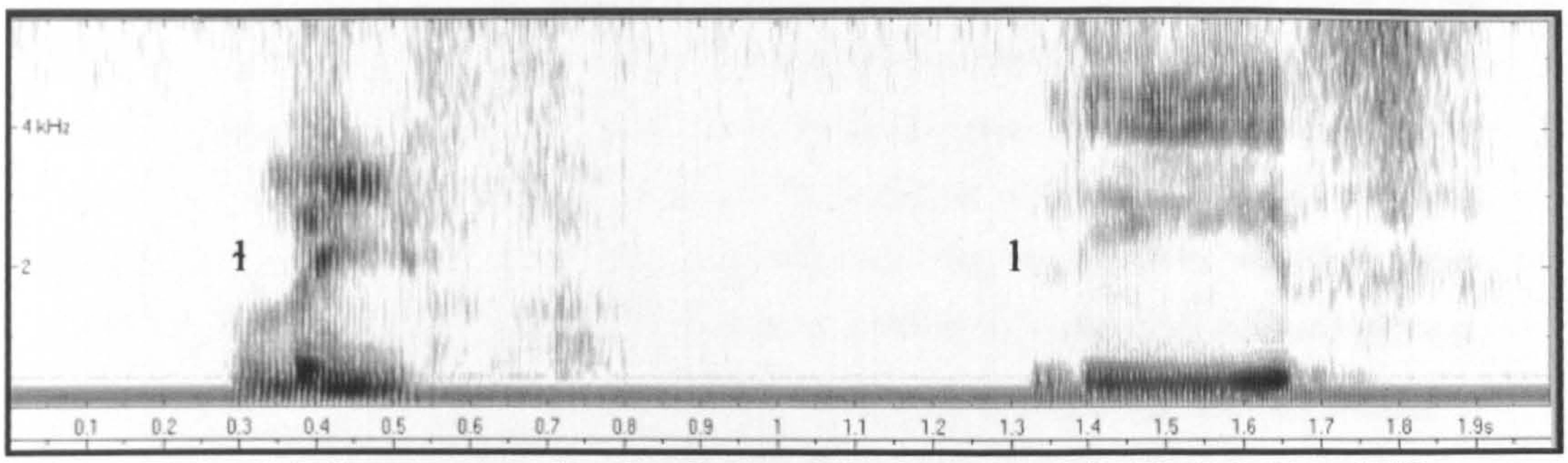


Figure 3.5: Spectrogram showing the word 'leaf' produced by EM7 (left) as [ɫif], F2 = 9.54Z; and by BM7 (right) as [li:f], F2 = 12.48Z.

3.7.2 Coda and syllabic position: adults

3.7.2.1 auditory analysis

Starting again with the data from the IViE corpus, Figure 3.6 shows the patterns found for /l/ in coda position. While most of the speakers produce the expected dark [ɫ] variant in this position, some speakers also produce some vocalised tokens. Vocalisations mainly occurred in tokens where /l/ was in coda (e.g. 'school'; 'ball') rather than syllabic position (e.g. 'people'; 'uncle'). There were also sporadic productions of clear [l]'s by 3 of the speakers, although clear [l]'s are not expected in this context. These occurred in the words 'meal', 'will', and 'pull' for F2, and 'meal' for F4 and F5.

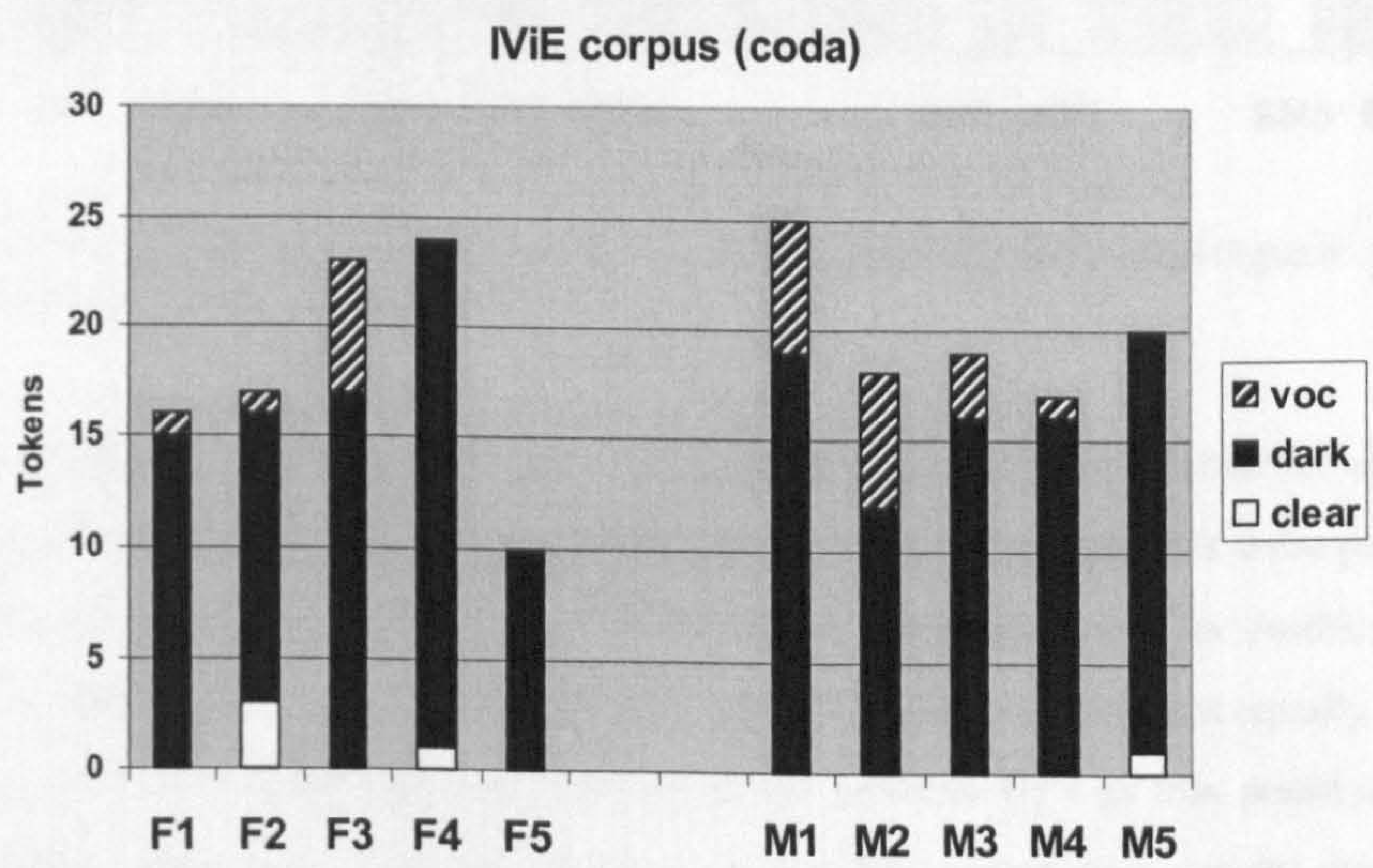


Figure 3.6: Results for syllable-coda and syllabic /l/ in English for the monolingual adults from the IViE corpus. N = 189.

Similar results are found for the monolinguals' parents from my study (Figure 3.7 and Table 3.5), as most of the realizations for /l/ fall in the dark category. It is interesting to note, however, that all six speakers also produce a small amount of vocalised /l/'s. As in the IViE data, vocalisations mainly occurred in tokens where /l/ was in coda rather than syllabic position. In fact, out of the 381 final /l/ tokens found for the six adults, a total of 70 were vocalised, with 63 out of possible 292 falling in coda position (22%) and only seven out of possible 88 in syllabic position (8%). The vocalisations varied between rounded and unrounded high back realisations and included mainly [ʊ], [ɤ], and [ɵ], e.g. 'ball' [bɔʊ], 'camel' [kamɵ], and 'bottle' [bɔtɤ]. A couple of realizations also seemed to be an extension of an offglide from a preceding vowel e.g. 'wall' [wɔ::], 'nail' [neɪᵊ]. Finally, all six monolingual subjects were consistent in the /l/ patterns they produced regardless of the style (reading list *versus* story telling and interview data) (Table 3.5).

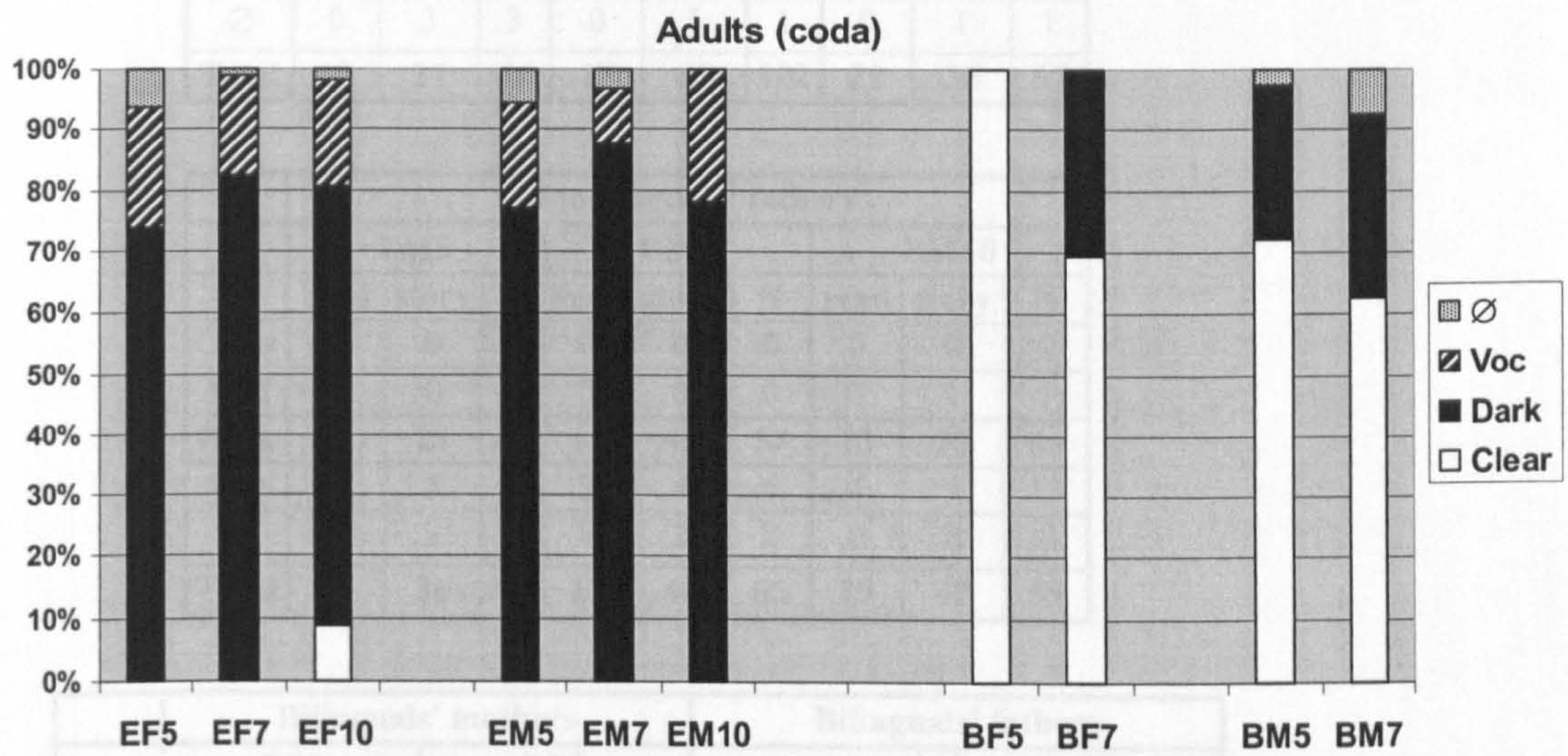


Figure 3.7: Results for coda and syllabic /l/ in English by the monolinguals' parents and the bilinguals' parents. N = 547

The bilinguals' parents, on the other hand, show clear evidence of their L1 affecting their production in a second language. A considerable number of coda /l/'s were produced as clear [l]'s by all four speakers, which is the pattern normally found for Arabic /l/'s in this context. The degree of L2 interference did not affect the four speakers equally, as can be seen by the fact that BF5 makes categorical use of clear [l]'s in this position while BM7 managed to produce 11 out of 16 tokens using the correct dark variety during the reading list activity (Table 3.5). Note, however, that BM7 did not manage to produce as many dark [ɫ]'s during the story telling activities and resorted to an almost categorical use clear [l]'s instead. Another aspect of language interference that showed in the parents' productions is the insertion of an epenthetic schwa before otherwise syllabic [l̩]'s following consonants in coda position e.g. 'kettle' [ˈkɛtəl]; 'marble' [ˈmɑːɾbəl]. Such a pattern is often mentioned as a feature of foreign-accented speech (Cruttenden, 2001: 160; Heselwood & McChrystal, 2000: 51). Moreover, none of the four subjects produced any /l/ vocalisations.

Table 3.5: Detailed results for /l/ production in coda and syllabic position during the reading list and story telling activities for the adults in English.

	Monolinguals' mothers								
	EF5			EF7			EF10		
	read	story	N	read	story	N	read	story	N
Clear	0	0	0	0	0	0	0	5	5
Med	0	1	1		0	0	0	4	4
Dark	13	20	33	14	70	84	17	20	37
Voc	6	3	9	3	14	17	4	6	10
Ø	0	3	3	0	1	1	0	1	1
Total	19	27	46	17	85	102	21	36	57

	Monolinguals' fathers								
	EM5			EM7			EM10		
	read	story	N	read	story	N	read	story	N
Clear	0	0	0	0	0	0	0	0	0
Med	0	0	0	0	0	0	0	1	1
Dark	12	28	40	15	42	57	15	30	45
Voc	4	5	9	2	4	6	4	9	13
Ø	0	3	3	0	2	2	0	0	0
Total	16	36	52	17	48	65	19	40	59

	Bilinguals' mothers						Bilinguals' fathers					
	BF5			BF7			BM5			BM7		
	read	story	N	read	story	N	read	story	N	read	story	N
Clear	18	10	28	6	30	36	14	16	30	5	20	25
Med	0	0	0	2	4	6	1	5	6	0	0	0
Dark	0	0	0	8	5	13	4	2	6	11	1	12
Ø	0	0	0	0	0	0	0	1	1	0	3	3
Total	18	10	28	16	39	55	19	24	43	16	24	40

3.7.2.2 acoustic analysis

Acoustic analysis was conducted on tokens produced in isolation in absolute word-final position. Figure 3.8 shows F2 distribution for final /l/ produced by the monolinguals' parents and the bilinguals' parents, with an indication of whether the token was heard as clear (white dashes), dark (black dashes) or vocalized (black circles). The bark-scaled figures correspond to measurements made from /l/ tokens following the vowels /iː/, /e/, /ɛ/, /a/, /ɔ/, /u/, and /ə/, as well as from tokens where /l/ was syllabic. As explained in Section 3.7.1, there were not enough tokens for each vocalic context to be presented separately, but an equal number of tokens from each context was included for each of the adults in order to obtain comparable data.

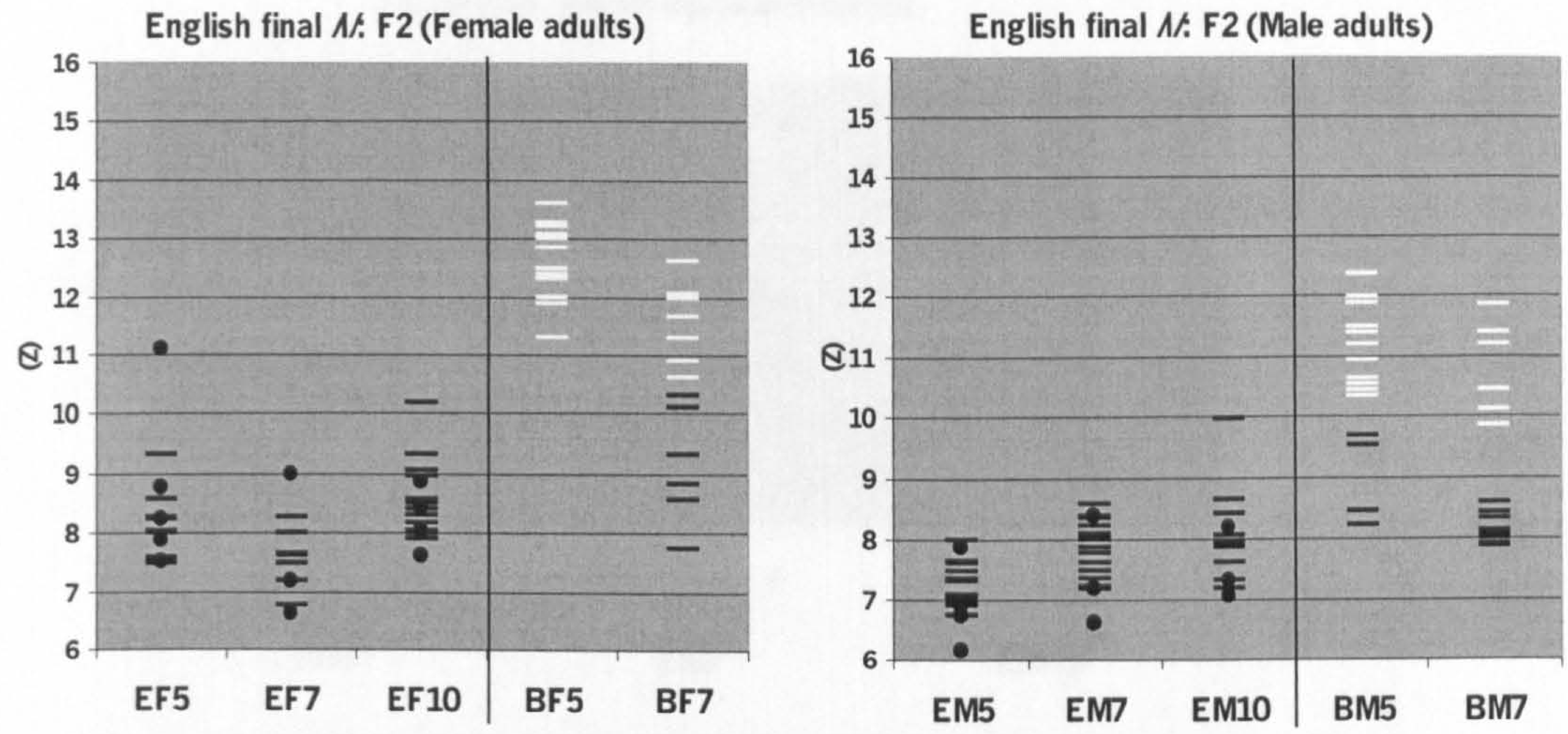


Figure 3.8: Bark-scaled F2 measurements for syllable-initial /l/ in English produced by the monolinguals' mothers (left) and fathers (right). White dashes indicate perceptually clear [l]'s, black dashes indicate perceptually dark [ɫ]'s, while black circles indicate perceptually vocalised /l/'s. N = 178.

Vocalized /l/'s on the whole have similar F2 frequencies to dark [ɫ]'s for the monolinguals' parents, reflecting a similar back quality. One exception is the word 'nail' by EF5, which was produced as [neɪə], and had an F2 of 11.1Z. Unlike the F2 distribution in syllable-initial position, there is hardly any overlap between the F2 measurements for clear and dark /l/'s in final position. Moreover, F2 measurements for final dark /l/'s are on the whole lower than those for initial dark [ɫ]'s (Figure 3.9), supporting the evidence that final dark [ɫ]'s in English are generally phonetically darker than initial dark [ɫ]'s (Carter, 1999). To find out whether the difference is significant, T-tests were run on the dark tokens found for EM5, EM7, and EM10 in initial and final position, since they are the three speakers that used initial dark [ɫ]'s the most, and therefore produced enough tokens to allow comparison with final position. Only the tokens with similar vocalic contexts following /l/'s in initial position and preceding /l/'s in final position were chosen for each speaker in order to obtain comparable data. The tests were highly significant for EM5 ($p = 0.001$) and significant for EM10 ($p = 0.01$) and EM7 and ($p = 0.02$). Figure 3.9 shows the distribution of initial and final dark /l/'s for the three speakers. Despite the overlap, the F2 ranges for initial and final position are clearly different, with the F2 frequencies in initial position being mainly concentrated between 8 and 10Z while in final position they are mainly concentrated between 7 and 8Z, never exceeding 9Z.

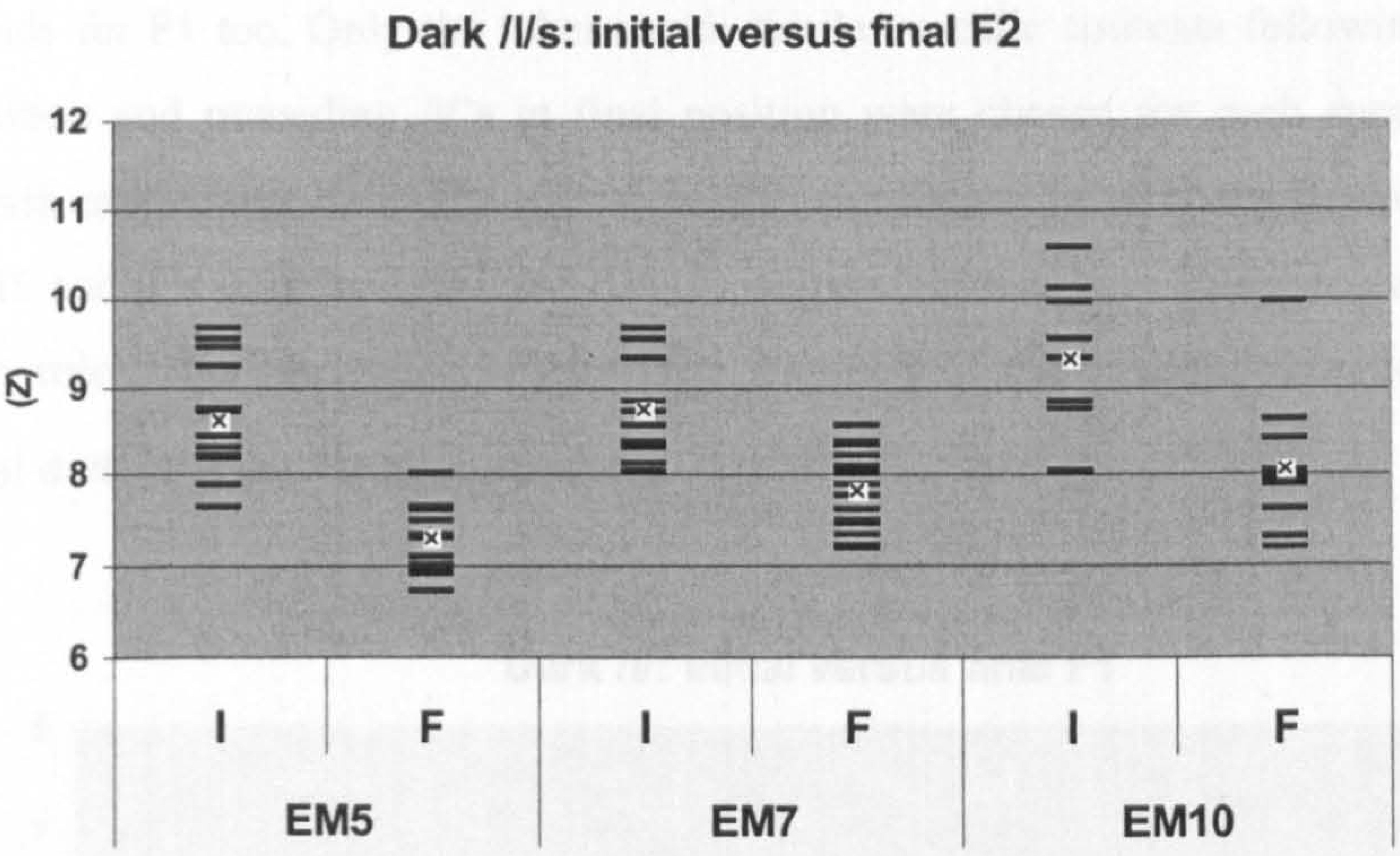


Figure 3.9: F2 distribution for initial and final dark [ɫ]'s produced by the three monolingual English fathers. Crosses indicate means. N = 40.

Moving on to F1, measurements show a tendency for dark [ɫ]'s to have a higher frequency than clear ones, but once again the difference is not as straightforward as for the F2 patterns and there is a high degree of overlap between clear and dark /l/'s with respect to F1 frequency (Figure 3.10).

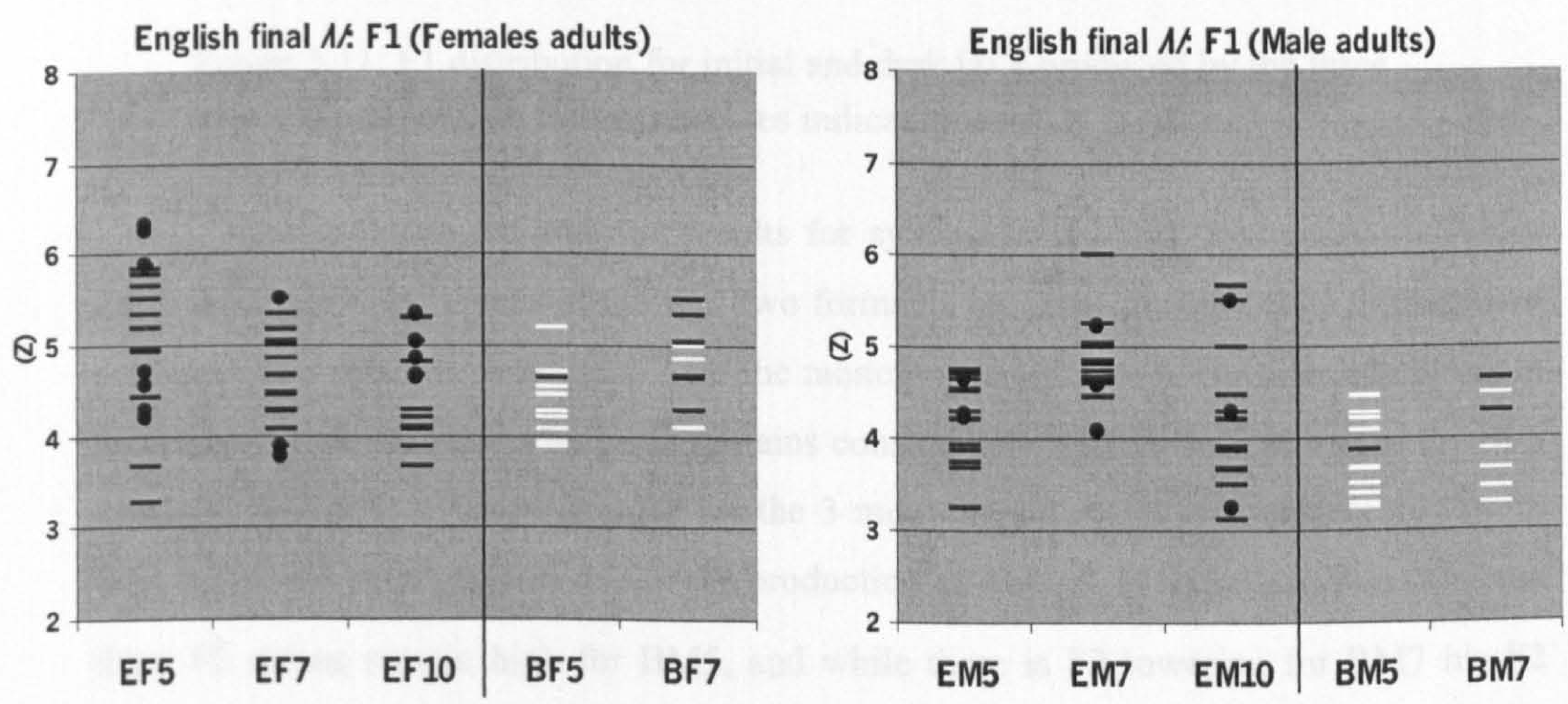


Figure 3.10: Bark-scaled F1 measurements for syllable-final /l/ in English produced by the monolinguals' mothers (left) and fathers (right). White dashes indicate perceptually clear [l]'s, black dashes indicate perceptually dark [ɫ]'s, while black circles indicate perceptually vocalised /l/'s. N = 178.

Since F2 measurements for final dark [ɫ]'s turned out to be lower than those for initial dark [ɫ]'s (Figure 3.9), T-tests were run on the F1 for dark tokens found for EM5, EM7, and EM10 in initial and final position in order to find out whether the significance stands for F1 too. Only the tokens with similar vocalic contexts following /l/'s in initial position and preceding /l/'s in final position were chosen for each speaker in order to obtain comparable data. The tests were also significant for all three speakers ($p = 0.02$ for EM5 and $p = 0.01$ for EM7 and EM10), suggesting that both F2 raising and F1 lowering play role in the perception of dark [ɫ]'s. Figure 3.11 shows the distribution of initial and final dark [ɫ]'s for the three speakers.

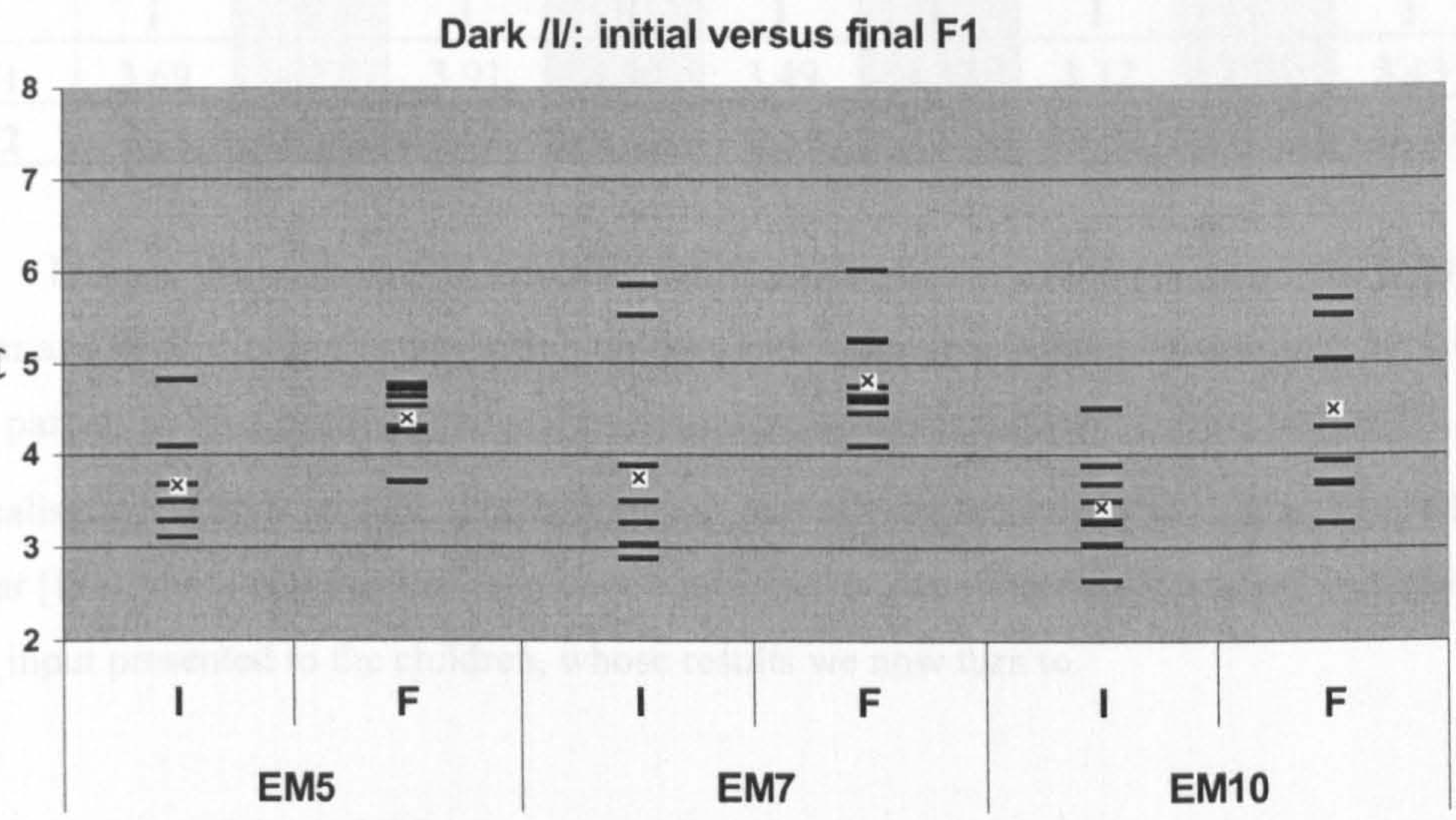


Figure 3.11: F1 distribution for initial and dark [ɫ]'s produced by the three monolingual English fathers. Crosses indicate means. N = 52.

To summarise the acoustic results for syllable-initial and final position, Tables 3.6 and 3.7 show the means of the first two formants obtained for the adults in these two positions. In Table 3.6, note that F2 for the monolingual females is considerably lower in final than initial position, whereas it remains considerably high in both positions for BF5 and BF7. In Table 3.7, note that F2 for the 3 monolingual males is considerably low in both initial and final position due to the production of dark [ɫ] in initial position, whereas there F2 means remain high for BM5, and while there is F2 lowering for BM7 his F2 mean in final position is almost similar to the highest F2 in initial position for the monolingual males.

Table 3.6: Mean F1 and F2 measurements for the monolinguals' mothers (left) and the bilinguals' mothers (right)

	Monolinguals' mothers						Bilinguals' mothers			
	EF5		EF7		EF10		BF5		BF7	
	I	F	I	F	I	F	I	F	I	F
F1	4.71	5.08	3.98	4.76	4.51	4.52	4.19	4.46	4.41	4.74
F2	11.70	7.99	10.66	7.56	10.81	8.61	12.48	12.78	12.29	10.35

Table 3.7: Mean F1 and F2 measurements for the monolinguals' fathers (left) and bilinguals' fathers (right)

	Monolinguals' fathers						Bilinguals' fathers			
	EM5		EM7		EM10		BM5		BM7	
	I	F	I	F	I	F	I	F	I	F
F1	3.69	4.27	3.91	4.90	3.49	4.37	3.32	3.74	3.43	3.80
F2	8.73	7.20	8.73	7.82	9.53	8.11	10.32	10.78	11.29	9.15

In sum, the adult targets available for /l/ production in initial position vary between clear and dark allophones depending on the gender, the geographical origin, and the L1 of the parent. In final position, the main variant for the monolinguals' parents is dark [ɫ], but vocalisation is also present. The bilinguals' parents, on the other hand, mainly produce clear [l]'s. These patterns indicate, once again, the degree of variability that is available in the input presented to the children, whose results we now turn to.

3.7.3 Onset position: children

3.7.3.1 Auditory analysis: picture naming and story telling activities

Results for the children are presented in raw figures rather than percentages, as some of the children produced a small number of tokens for a given context, and percentages might therefore be misleading. Figure 3.12 shows /l/ patterns for the monolingual and bilingual children in syllable-onset position. Starting with the monolingual children first, it is surprising that they produce very few initial dark /l/'s, even those whose parents were reported as dark [ɫ] users in Figure 3.2 (e.g. E7's parents). The production of medium and dark [ɫ] by each child ranges between 2 and 6 tokens only (Table 3.8), and these are mainly contextualised in that they are only produced in the environment of back vowels (e.g. 'La La'). The bilingual children behave very similarly to the monolingual children and mainly produce clear [l]'s.

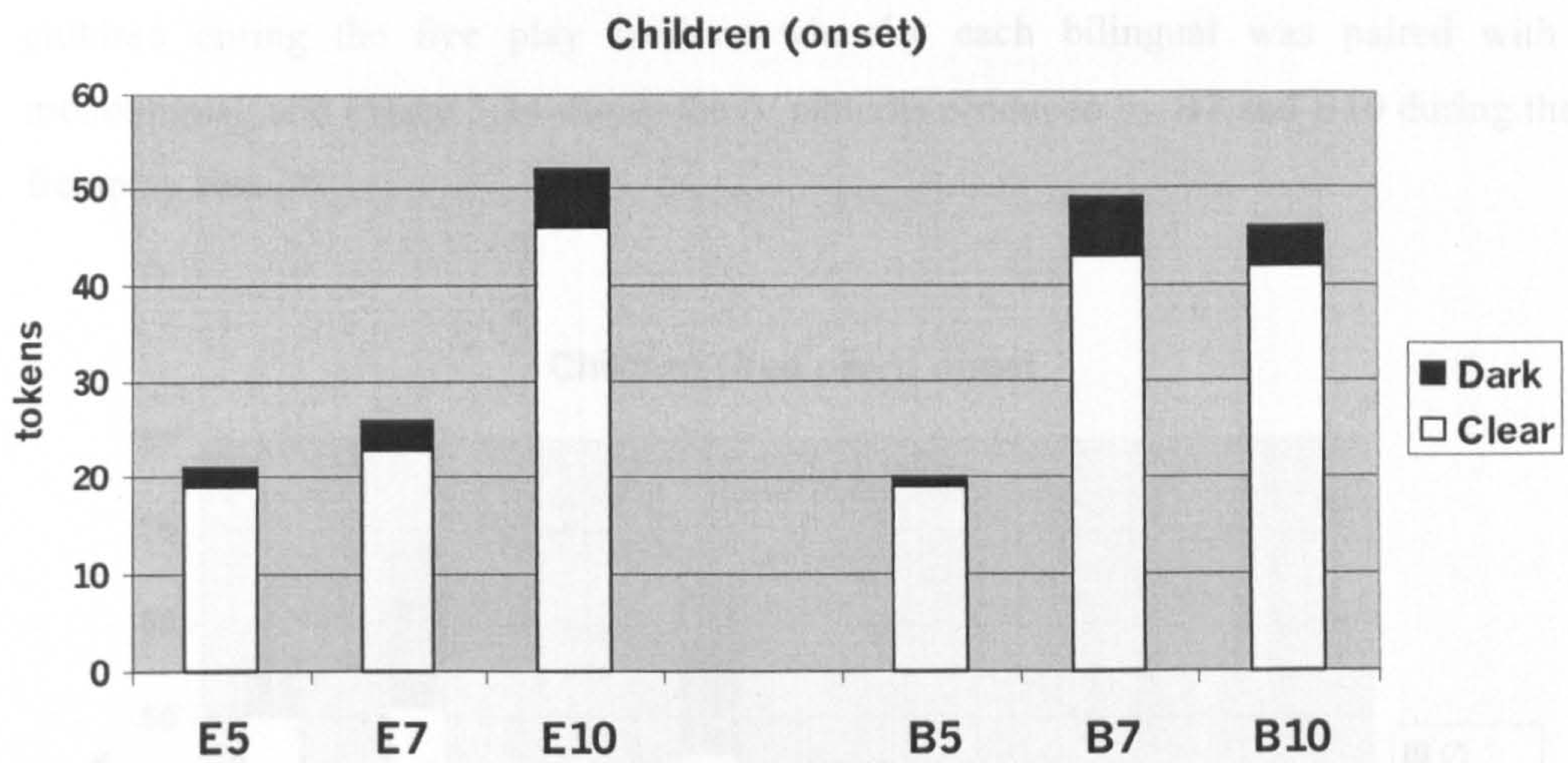


Figure 3.12: Results for /l/ in onset position in English for the monolingual and bilingual children. N = 214.

Table 3.7: Detailed results for /l/ production in syllable-onset position during the picture naming and story telling activities for the children in English.

	Monolingual children								
	E5			E7			E10		
	pic	story	N	pic	story	N	pic	story	N
Clear	10	9	19	10	13	**23**	14	32	46
Med	0	1	1	0	1	**1**	0	2	2
Dark	1	0	1	2	0	**2**	1	3	4
Total	**11**	**10**	**21**	**12**	**14**	**26**	**15**	**37**	**52**

	Bilingual children								
	B5			B7			B10		
	pic	story	N	pic	story	N	pic	story	N
Clear	7	12	19	11	32	43	15	27	42
Med	0	0	0	0	5	5	0	3	3
Dark	1	0	1	1	0	1	1	0	1
Total	**8**	**12**	**20**	**12**	**37**	**49**	**16**	**30**	**46**

3.7.3.2 Auditory analysis: free-play sessions

Since I conducted the picture-naming and story telling activities, material from the free-play sessions between the children was used in order to support the findings reported in Section 3.7.4.1. As mentioned in Chapter Two, each of the bilingual children was recorded playing with a monolingual friend of the same age and B7 and B10 were also recorded playing together in order to test any possible difference in the bilinguals' linguistic behaviour depending on whether they are interacting with monolinguals or bilinguals. Figure 3.13 and Table 3.8 show the /l/ patterns produced by each of the 6 children during the free play sessions whereby each bilingual was paired with a monolingual, and Figure 3.14 shows the /l/ patterns produced by B7 and B10 during their free-play session.

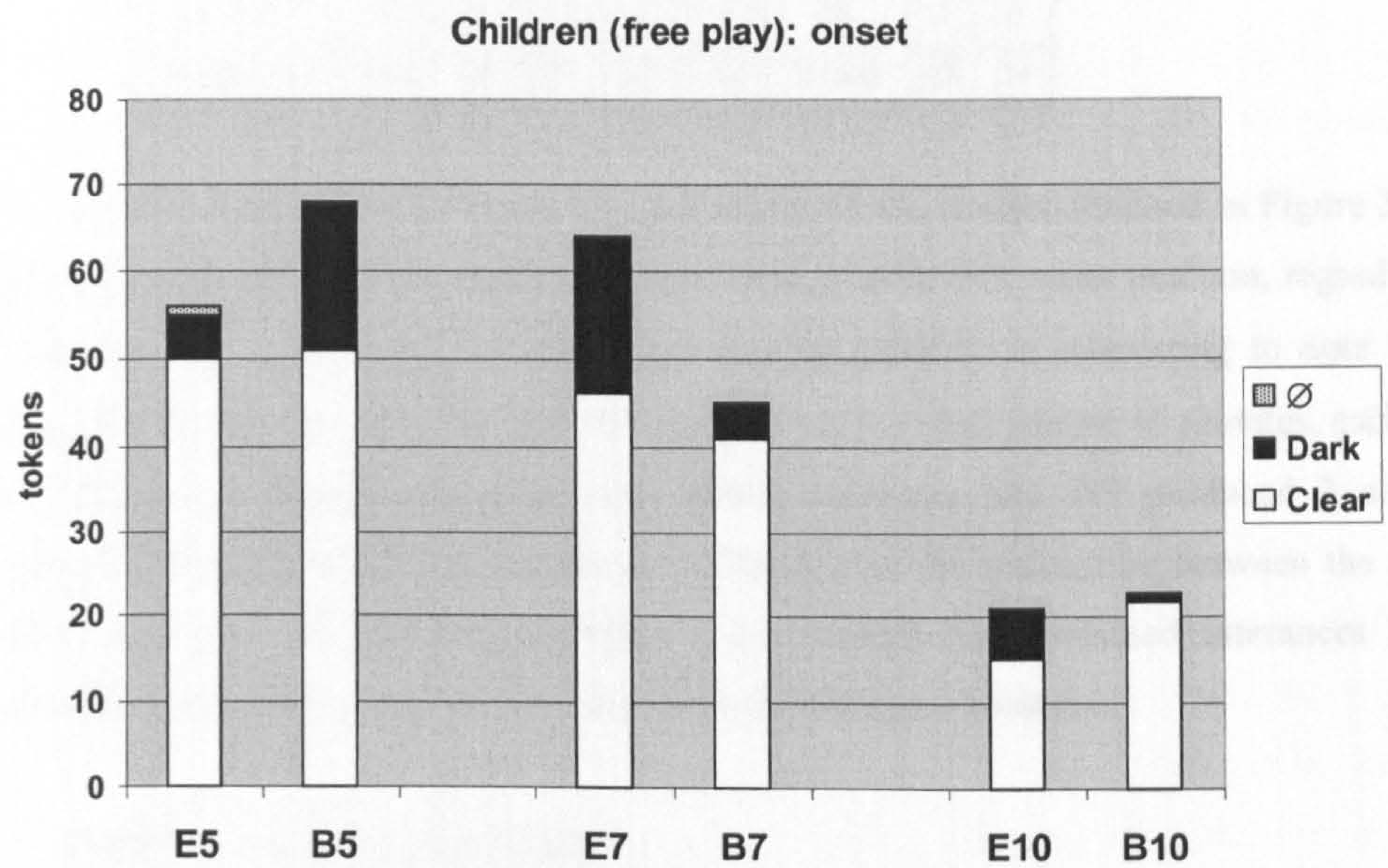


Figure 3.13: Results for the /l/ patterns in onset position that were found during the paired free-play sessions between the monolingual and the bilingual children. N = 277.

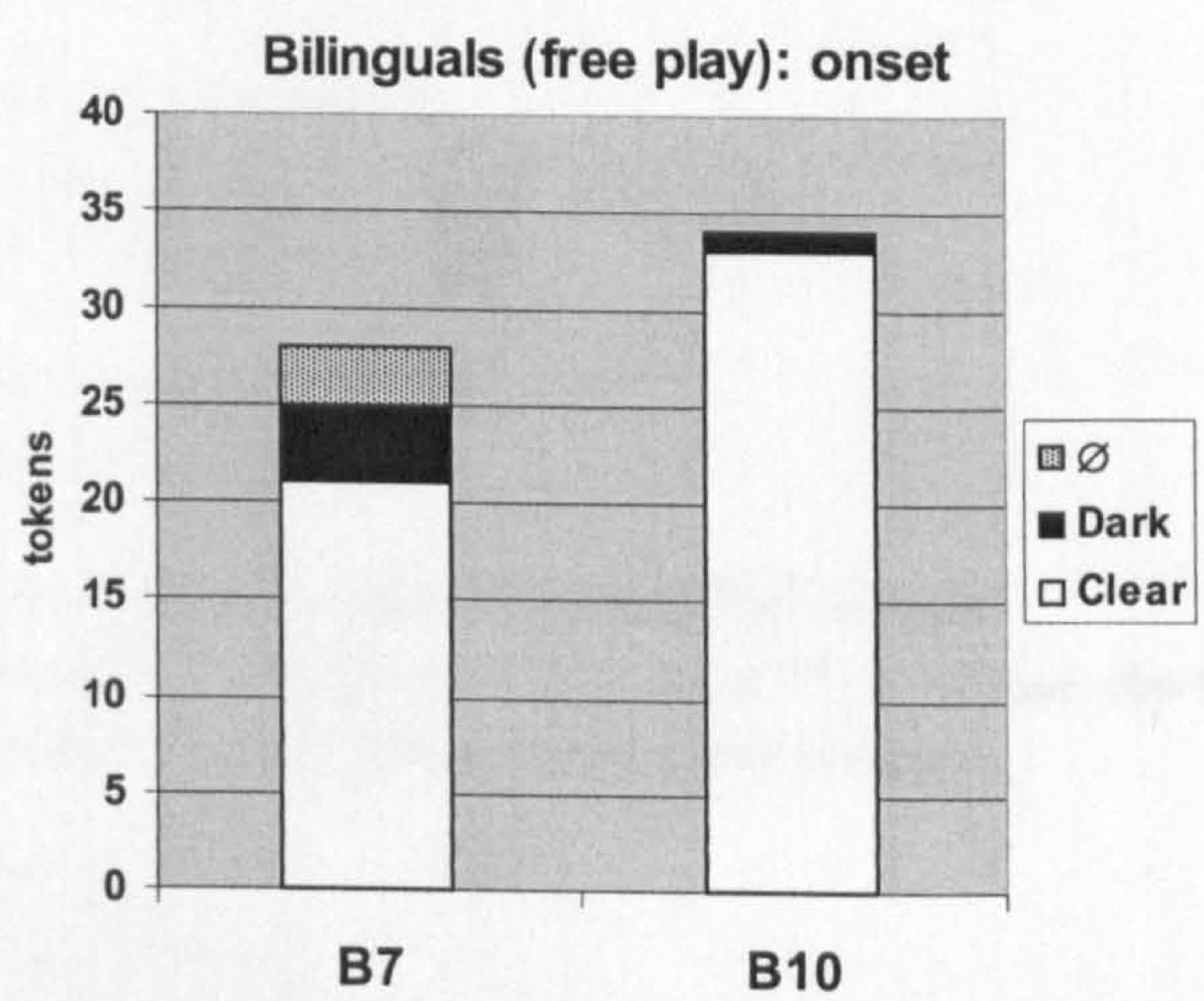


Figure 3.14: Results for the /l/ pattern in onset position that were found during the paired free-play sessions between two of the bilingual children. N = 62.

Table 3.8: Detailed results for /l/ pattern in onset position during the paired free-play sessions between the children.

Bilingual + Monolingual							B7 + B10		
Onset	**E5**	**B5**	**E7**	**B7**	**E10**	**B10**	**Onset**	**B7**	**B10**
Clear	50	51	46	41	15	22	**Clear**	21	33
Med	4	7	9	3	5	1	**Med**	4	1
Dark	1	10	9	1	1	0	**Dark**	0	0
Ø	1	0	0	0	0	0	**Ø**	3	0
Total	56	68	64	45	21	23	**Total**	28	34

Results from Figure 3.13 and 3.14 are similar to the results obtained in Figure 3.12 and support the fact that the children mainly used clear [l]'s in onset position, regardless of whether the interlocutor was myself or another child. It is interesting to note that during the free-play session between B7 and B10 which lasted around 45 minutes, each of the bilinguals produced only one short Arabic utterance and B7 produced 3 code-switched utterances while B10 produced 5. The rest of the interaction between the two brother was only in English. Examples 1 to 6 show code-switched utterances that contained /l/ tokens and that were produced by the bilingual brothers:

(1) B7: *The tape's* fɑːdˤe kɪlːa
The tape's empty all-it (3rd pers.fem.)
The tape's empty, all of it.

(2) B7: ʔane jɛlo
me yellow
I want the yellow one

(3) B7: laʔ baʕd ma xolˤɪsˤ
No yet not finish-past-3rd pers.masc.
No, it's not finished yet

(4) B10: jalːa *come on*
come on come on

(5) B10: *Go and tell* [tɛɫ] *Ghada* ʔɪnːo xolˤɪsˤ ɪl tape
Go and tell Ghada that finish-past-3rd pers.masc. the tape
Go and tell Ghada that the tape has finished.

(6) B10: laʔ ʔane jɛlo
no me yellow
No, I want the yellow one

Not many observations can be made from this small number of utterances, but the examples do show that B7 and B10 seem to be producing language-appropriate /l/ variants in the Arabic and mixed utterances, though the production of the word 'yellow' by both children sounded more Arab-like, especially with the use of the back close vowel [o].

3.7.4 Syllable-coda and syllabic position: children

3.7.4.1 Auditory analysis: picture-naming and story telling

Figure 3.15 shows the patterns found for the monolingual (left) and bilingual (right) children. Apart from the tokens presented in this figure, each of E5, B5 and B10 produced one or two tokens where /l/ was deleted; these are presented in Table 3.9. The most striking result is that the two groups and all six children display similar patterns with respect to /l/ production in this context. None of the bilinguals behaves like their parents by either producing a majority of clear [l]'s in this context or inserting a schwa before syllabic /l/'s. Syllable-final /l/'s were mainly realized as dark or vocalized during the picture naming activities (Table 3.9), but there was a small number of clear [l]'s that were produced by the bilinguals during the story telling activities. There are two ways of interpreting the occurrence of clear [l]'s in the bilinguals' production in the latter type of activity. First, knowing that one of the bilinguals' fathers, BM7, also managed to show a greater ability to produce dark [ɫ]'s in words in isolation but not during the story telling activities, the influence from Arabic might therefore be more easily detectable in the bilinguals' production in running speech rather than single word elicitation. Since B5 produces more clear [l]'s in this context than B7 and B10, this could be interpreted as a increasing ability for the bilinguals to keep the variants they choose for each language separate. On the other hand, B5's behaviour might be showing a developmental feature in her speech, since it was mentioned in Section 3.1.4 that dark [ɫ]'s are acquired later than clear [l]'s by children, and even E5 produces the occasional clear [l] (Table 3.9).

It is interesting to note that the amount of vocalisation by the children is greater than that found for the monolinguals' parents in Figure 3.4, and it seems to gradually increase rather than decrease with age. This rules out the possible interpretation that vocalisation is simply a developmental feature that gradually disappears with age (e.g. Cruttenden, 2001). /l/ vocalisation is maintained by the children and, since it was also found in the production of the monolinguals' parents, it seems to be an established feature of the local accent. Of course, more research is needed to support this claim, but this is beyond the scope of the current study.

The types of vocalizations used by the children were similar to those found for the monolinguals' parents in Figure 3.6, but also extended to more open vowels like [ɐ] and [ʌʊ] (e.g. 'feel' [fiɐ]; 'table' [ˈteɪbʌʊ]). While the older subjects (E10 and B10) produced fewer vocalisations in syllabic than in coda position, thus similar to the adults, Table 3.10 shows that there was no clear pattern in the younger children's productions.

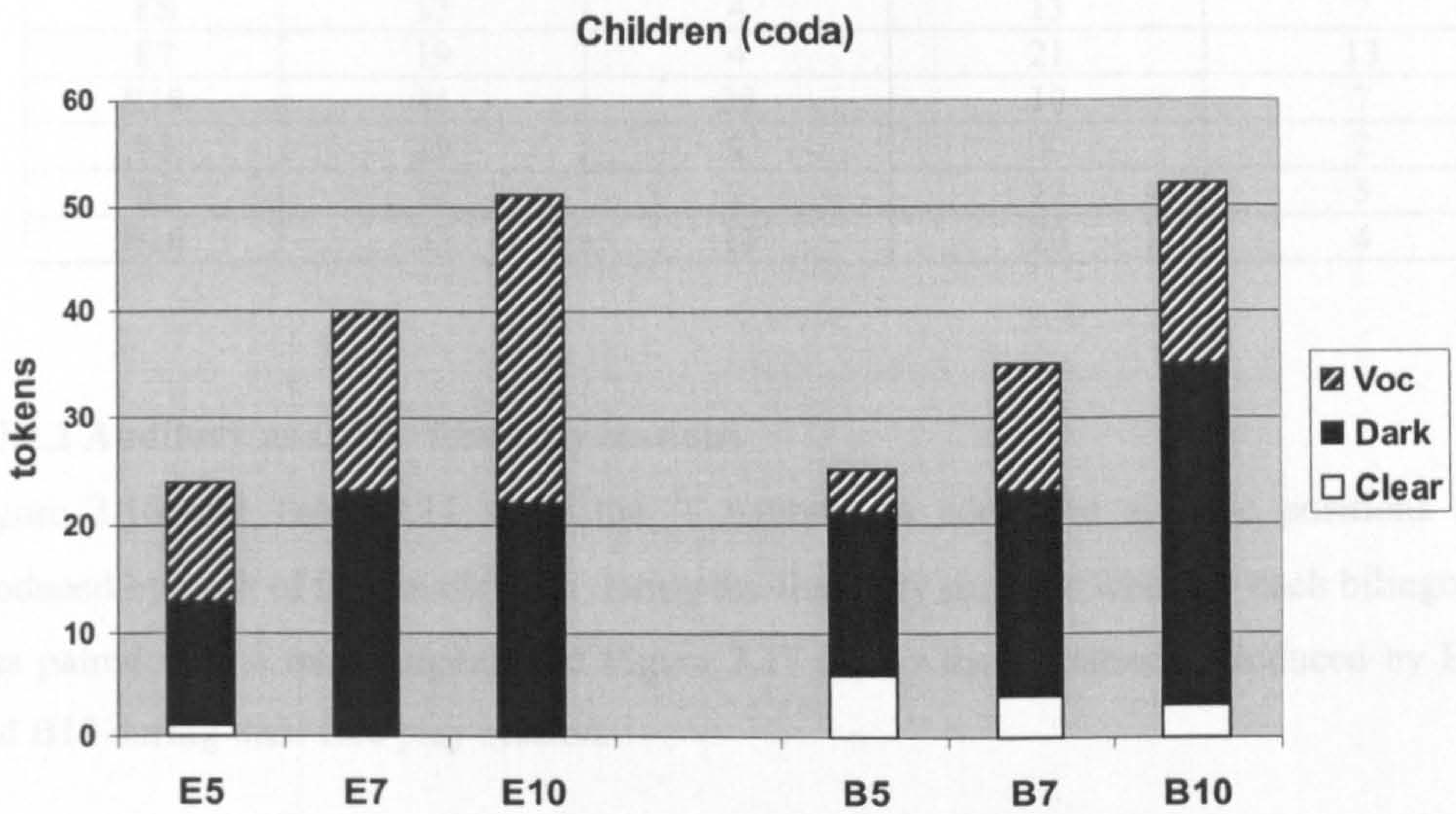


Figure 3.15: Results for syllable-final /l/ in English by the monolingual and bilingual children. E = English; B = Bilingual. N = 229.

Table 3.9: Results for /l/ production in syllable-coda and syllabic position during the picture naming and story telling activities for the children in English.

	Monolingual children								
	E5			E7			E10		
	pic	story	N	pic	story	N	pic	story	N
Clear	0	1	1	0	0	0	0	0	0
Dark	9	3	12	9	14	23	6	16	22
Voc	9	2	11	8	9	17	12	17	29
Ø	2	0	2	0	0	0	0	0	0
Total	**20**	**6**	**26**	**17**	**23**	**40**	**18**	**33**	**51**

	Bilingual children								
	B5			B7			B10		
	pic	story	N	pic	story	N	pic	story	N
Clear	0	6	6	2	2	4	0	3	3
Dark	11	4	15	8	11	19	8	24	32
Voc	4	0	4	8	4	12	13	4	17
Ø	2	0	2	0	0	0	0	1	1
Total	**17**	**10**	**27**	**18**	**17**	**35**	**21**	**1**	**53**

Table 3.10: Vocalised /l/ tokens in syllable-coda and syllabic position produced by the children.

	Coda position		Syllabic position	
	N	Vocalised	N	Vocalised
E5	13	4	13	7
E7	19	4	21	13
E10	41	22	10	7
B5	19	2	8	2
B7	23	7	12	5
B10	33	13	20	4

3.7.4.2 Auditory analysis: free-play sessions

Figure 3.16 and Table 3.11 show the /l/ patterns in coda and syllabic positions as produced by each of the six children during the free-play sessions whereby each bilingual was paired with a monolingual, and Figure 3.17 shows the /l/ patterns produced by B7 and B10 during their free play session.

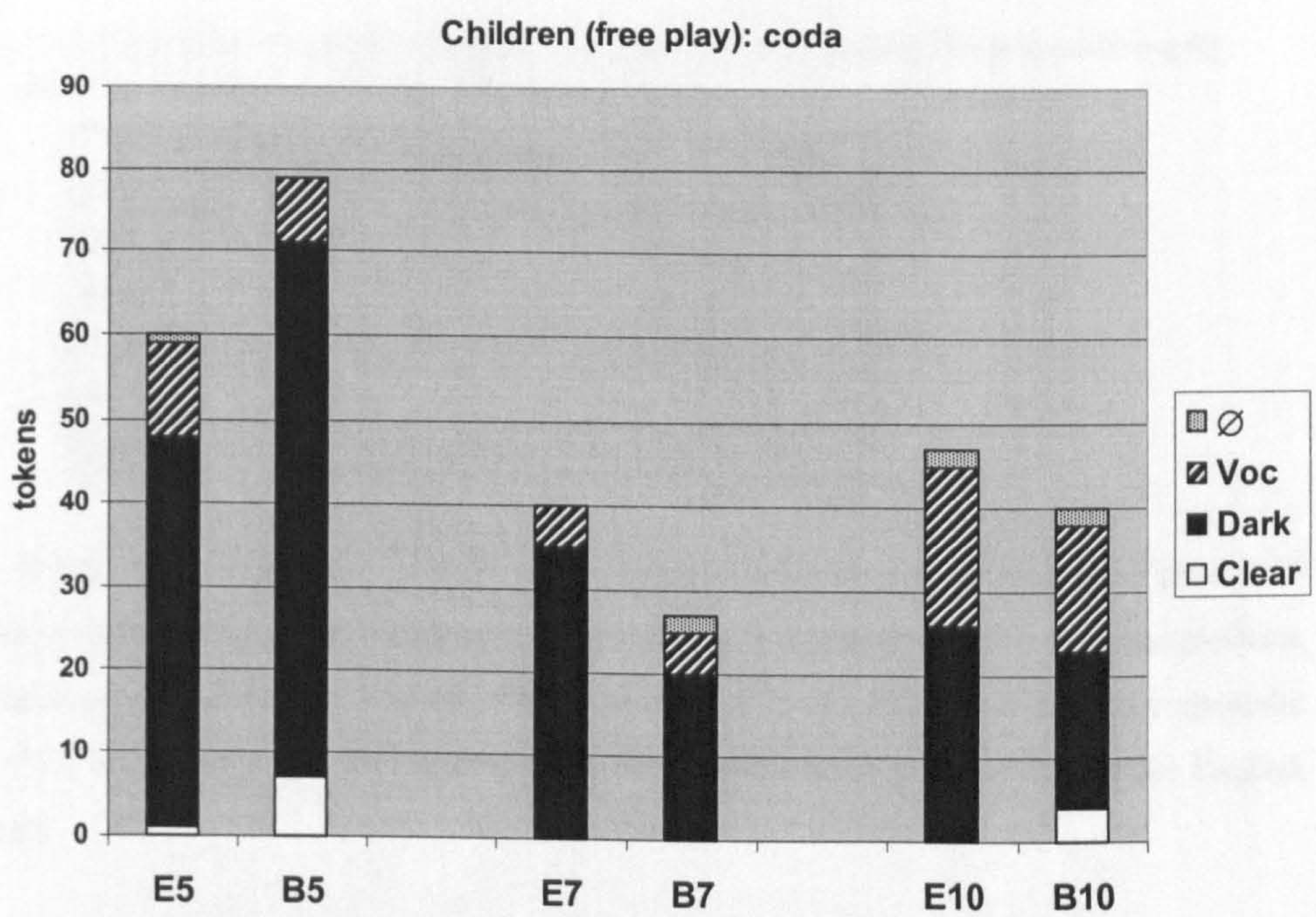


Figure 3.16: Results for the /l/ pattern in coda and syllabic position that were found during the paired free-play sessions between the monolingual and the bilingual children. N = 293.

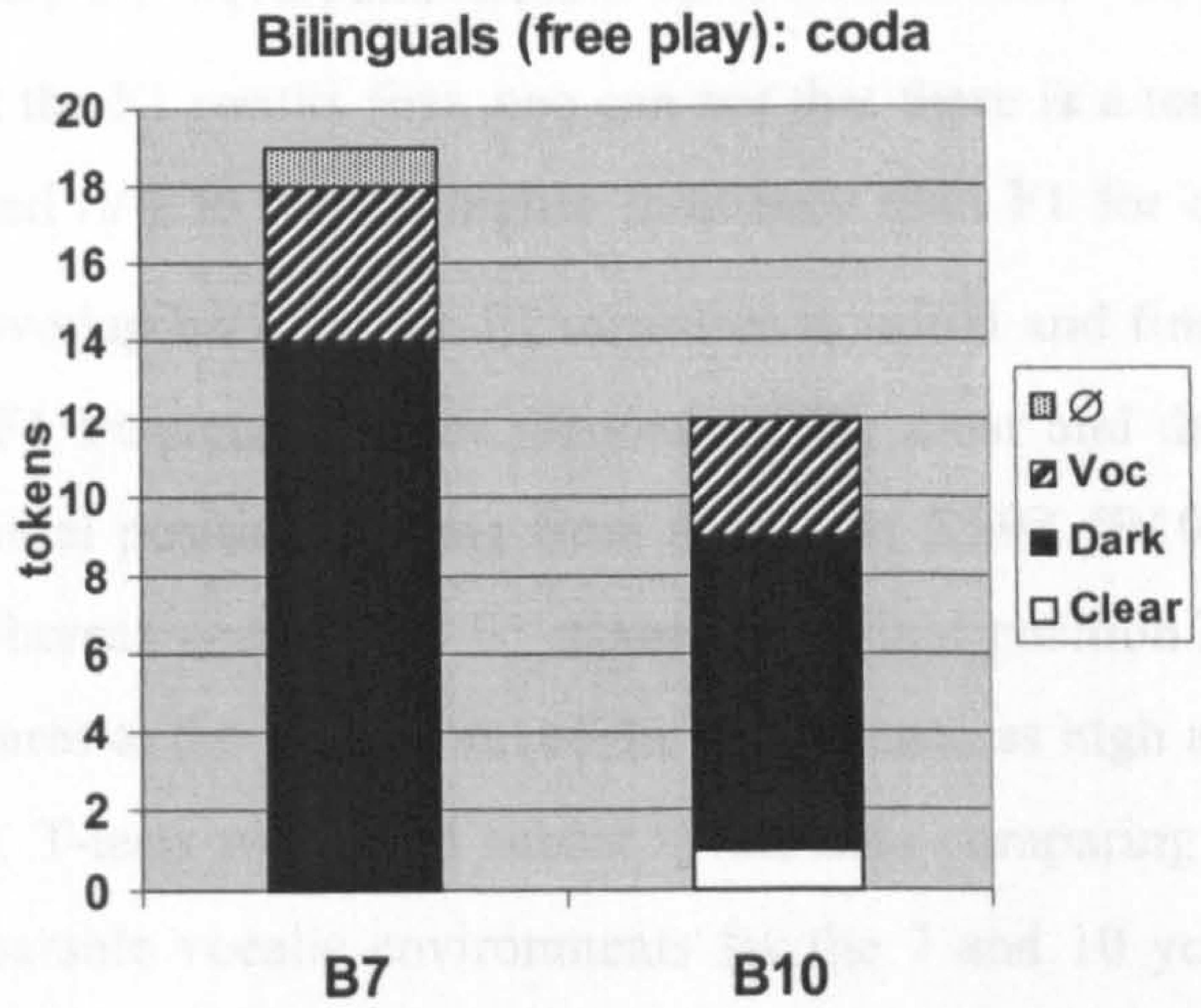


Figure 3.17: Results for the /l/ pattern in coda and syllabic position that were found during the paired free-play sessions between B7 and B10. N = 31.

Table 3.11: Detailed results for /l/ pattern in coda position during the paired free-play sessions between two of the bilingual children.

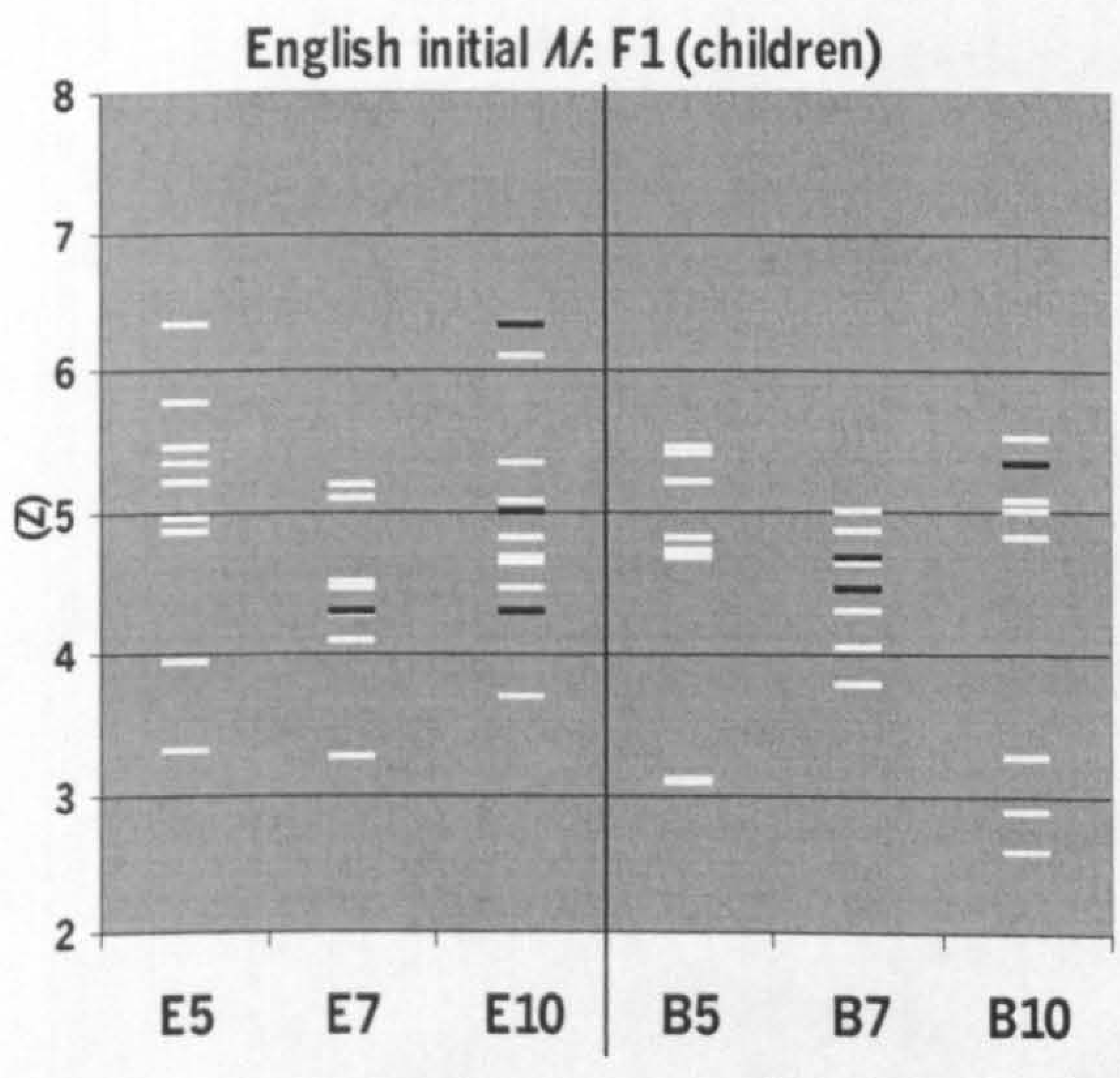
Bilingual + Monolingual							B7 + B10		
Coda	**E5**	**B5**	**E7**	**B7**	**E10**	**B10**	**Coda**	**B7**	**B10**
Clear	1	7	0	0	0	4	**Clear**	0	1
Dark	47	64	35	20	26	19	**Dark**	14	8
Voc	11	8	5	5	19	15	**Voc**	4	3
Ø	1	0	0	2	2	2	**Ø**	1	0
Total	**60**	**79**	**40**	**27**	**47**	**40**	**Total**	**19**	**12**

Results from Figures 3.16 and 3.17 are similar to the results obtained in Figure 3.15 and support the fact that the children mainly used dark or vocalized /l/'s in coda position, that vocalization seems to increase with age, and that the bilinguals produce sporadic clear [l]'s in running speech, even during their interactions with monolingual English children.

3.7.4.3 Acoustic analysis

Acoustic analysis was conducted on tokens in absolute word-initial and word-final position from words produced in isolation during the picture-naming activities. Figures 3.18 and 3.19 shows F1 and F2 distribution for initial and final /l/'s produced by the monolingual and bilingual children, with an indication of whether the token was heard as clear, dark, or vocalized. In initial position, the bark-scaled figures correspond to measurements made from /l/ tokens followed by the same vowels that were chosen for the adults in this position, mainly /iː/, /eɪ/, /a/, /aɪ/, /ɔ/, or /u/. Similarly, in final position the bark-scaled figures correspond to measurements made from /l/ tokens following the vowels /iː/, /eɪ/, /ɛ/, /a/, /ɔ/, /u/, and /ə/, as well as from tokens where /l/ was syllabic.

Looking at the F1 results first, one can see that there is a tendency for F1 for dark [ɫ]'s and vocalized /l/'s to have a higher frequency than F1 for clear [l]'s. Despite the great degree of overlap between the F1 measures in initial and final position, there seem to be exclusive F1 frequency ranges for each of the clear and dark variants in that F1 frequencies in initial position starting from as low as 2.59Z (B10) and never extending beyond 6.32Z, whereas none of the F1 measures in final position has a frequency below 4Z, but the measures at the upper level of the range reach as high as 7.87Z, especially for vocalized tokens. T-tests were on a subset of the data comparing F1 in initial and final position in comparable vocalic environments for the 7 and 10 year-olds only (the five-year-olds ended up with too few tokens when the vocalic context was contolled), and only E7 and E10 had a significant difference between their F1 in initial and final position ($p < 0.01$).



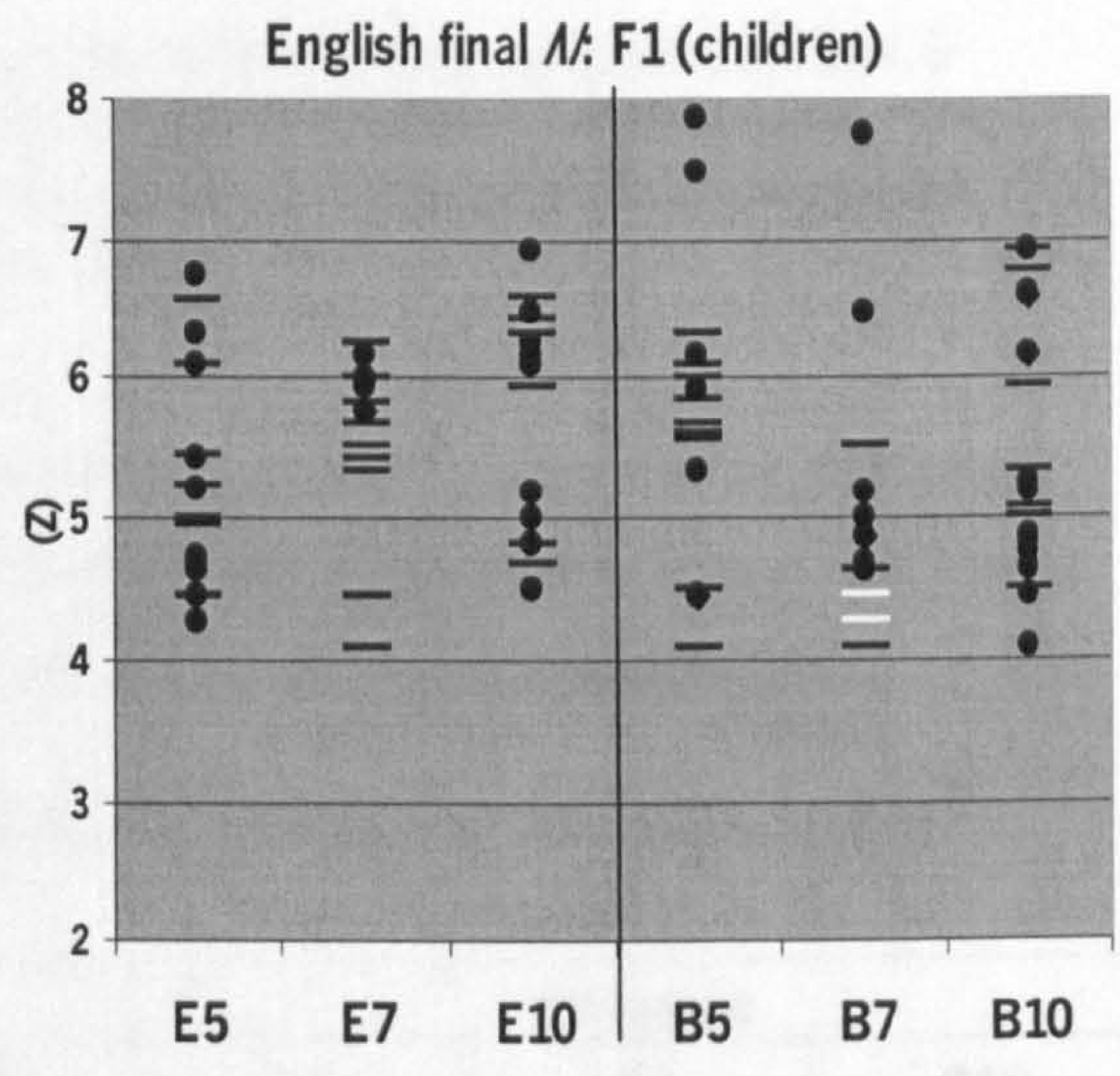


Figure 3.18: Bark-scaled F1 measurements for syllable-initial and syllable-final /l/ in English produced by the monolingual and bilingual children. White dashes indicate perceptually clear [l]'s, while black dashes indicate perceptually dark [ɫ]'s, and black circles indicate vocalic /l/'s. N = 105

As for F2, the difference in measurements between initial and final position is greater than that for F1, and the overlap is smaller. While F2 frequency for clear [l]'s in initial position mainly ranges between 11 and 15Z, in final position it is mainly concentrated in the 8 to 11Z region, therefore constituting an almost separate range of frequencies from that found for initial position. As for the slightly higher frequencies for some of the vocalized /l/'s (especially for E10 and B10), these resulted either from open-type vocalizations that were produced by the children such as [ˈfiɐ] for 'feel'; [tʰeɪɐ] for 'tail', or central schwa-like realization such as [neɪə] for 'nail' and [ˈtaʊə] for 'towel'.

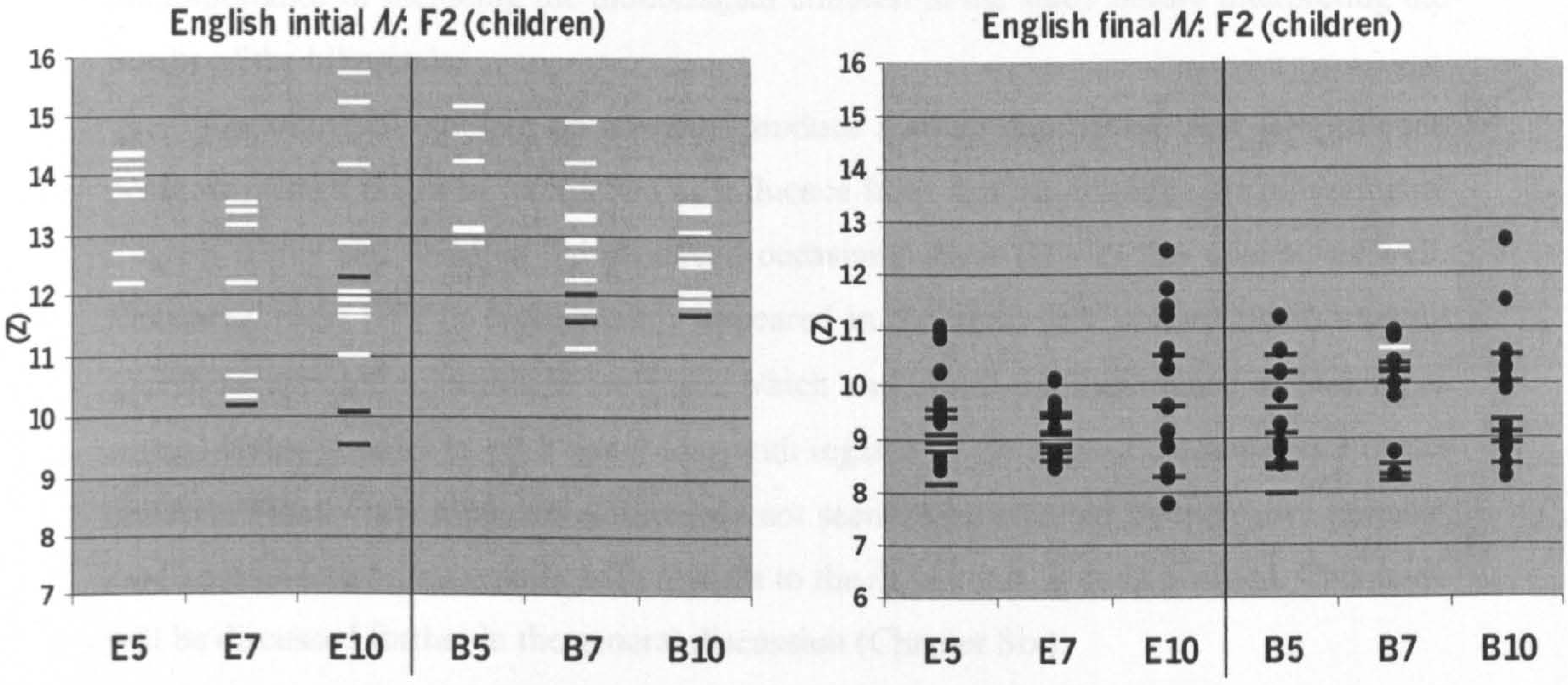


Figure 3.19: Bark-scaled F2 measurements for syllable-initial and syllable-final /l/ in English produced by the monolingual and bilingual children. White dashes indicate perceptually clear [l]'s, while black dashes indicate perceptually dark [ɫ]'s, and black circles indicate vocalic /l/'s. N = 105.

To summarise the acoustic results for syllable-initial and final position for the children, Table 3.12 shows the means of F1 and F2 obtained in these two contexts. Note how F2 is significantly lower in final position as opposed to initial position for all six children.

Table 3.12: Mean F1, F2 and F3 measurements for the monolingual (left) and bilingual children (right).

	Monolingual						Bilingual					
	E5		E7		E10		B5		B7		B10	
	I	F	I	F	I	F	I	F	I	F	I	F
F1	5.19	5.30	4.36	5.27	4.88	5.73	4.39	5.61	4.46	4.92	4.55	5.59
F2	13.94	9.52	12.13	9.04	12.36	9.95	13.72	9.48	13.19	10.23	12.52	9.64

3.7.5 Summary of the English results

In Figures 3.20 and 3.21, the results for children from each language group have been clustered together and presented along with those of the adults for an overall comparison. The most striking result remains that both the monolingual English children and the bilingual children have similar production patterns; these patterns show influence from the adult patterns but are also influenced by each of the children's age and linguistic background. Both groups of children have acquired the basic clear/dark /l/ distinction between onset and coda /l/'s, along with vocalisation. Though dark initial [ɫ]'s are also part of the accent of the English adults in this study, the children do not seem to be following this pattern and are opting for more clear [l]'s instead. Such findings underline the importance of including the monolingual children in the study before interpreting the results of the bilinguals.

The bilingual children do however produce a small number of clear [l]'s in coda positions, which might be interpreted as influence from Arabic, although the monolingual English adults and children did produced occasional clear [l]'s in this context as well. Moreover, clear [l]'s in codas mainly appeared in the children's production in running speech as opposed to words in isolation, which underlines the importance of looking at several styles in order to get a better idea with regards to the linguistic competence of the children. Finally, the bilingual children do not seem to be affected by their own parents' L2-like productions, especially with regards to the /l/ patterns in coda position. This issue will be discussed further in the general discussion (Chapter Six).

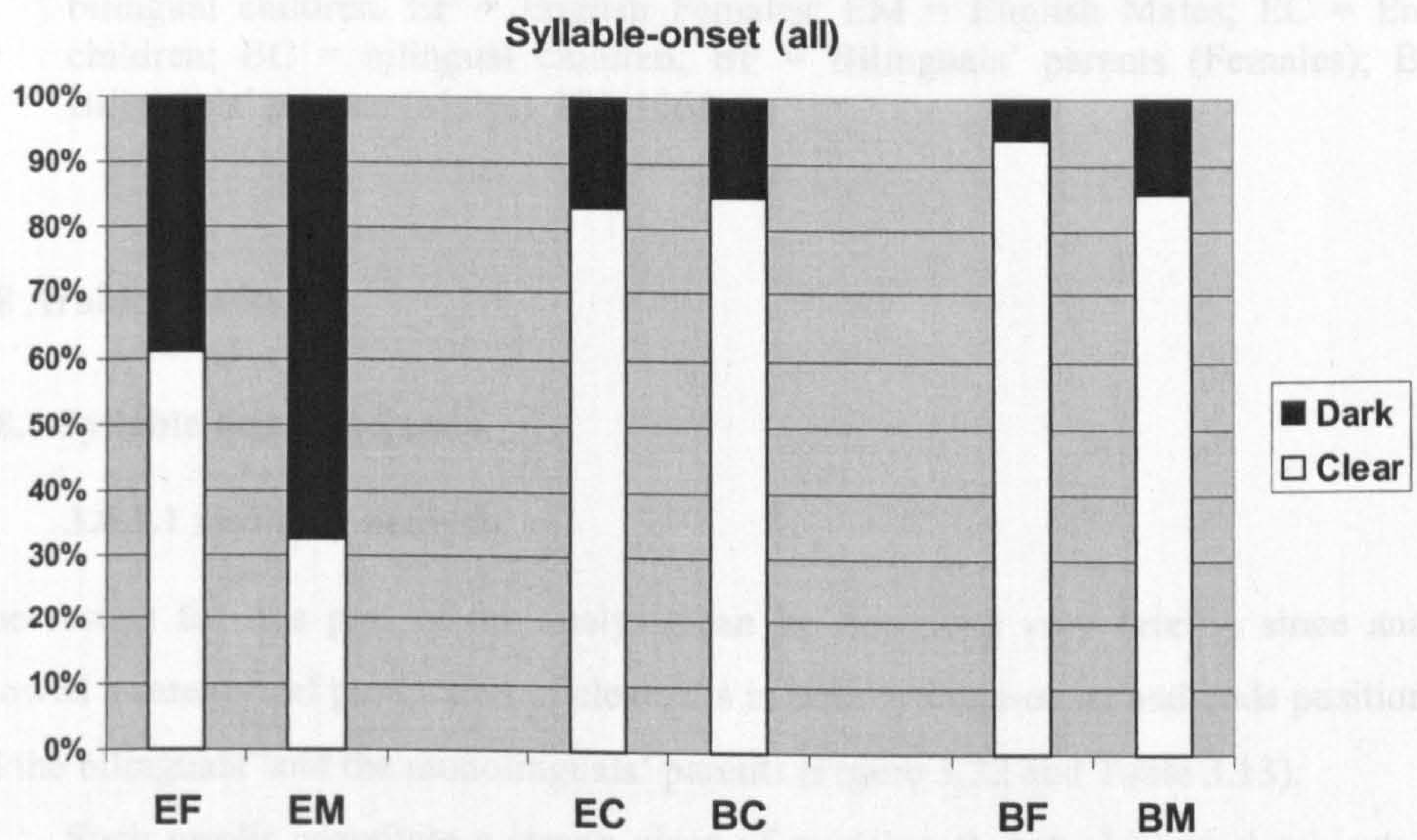


Figure 3.20: Results for syllable-onset /l/ in English by the monolinguals' parents, the bilinguals' parents, and the monolingual English and bilingual children. EF = English Females; EM = English Males; EC = English children; BC = bilingual children; BF = Bilinguals' parents (Females); BM = Bilinguals' parents (Males). N (tokens) = 1145.

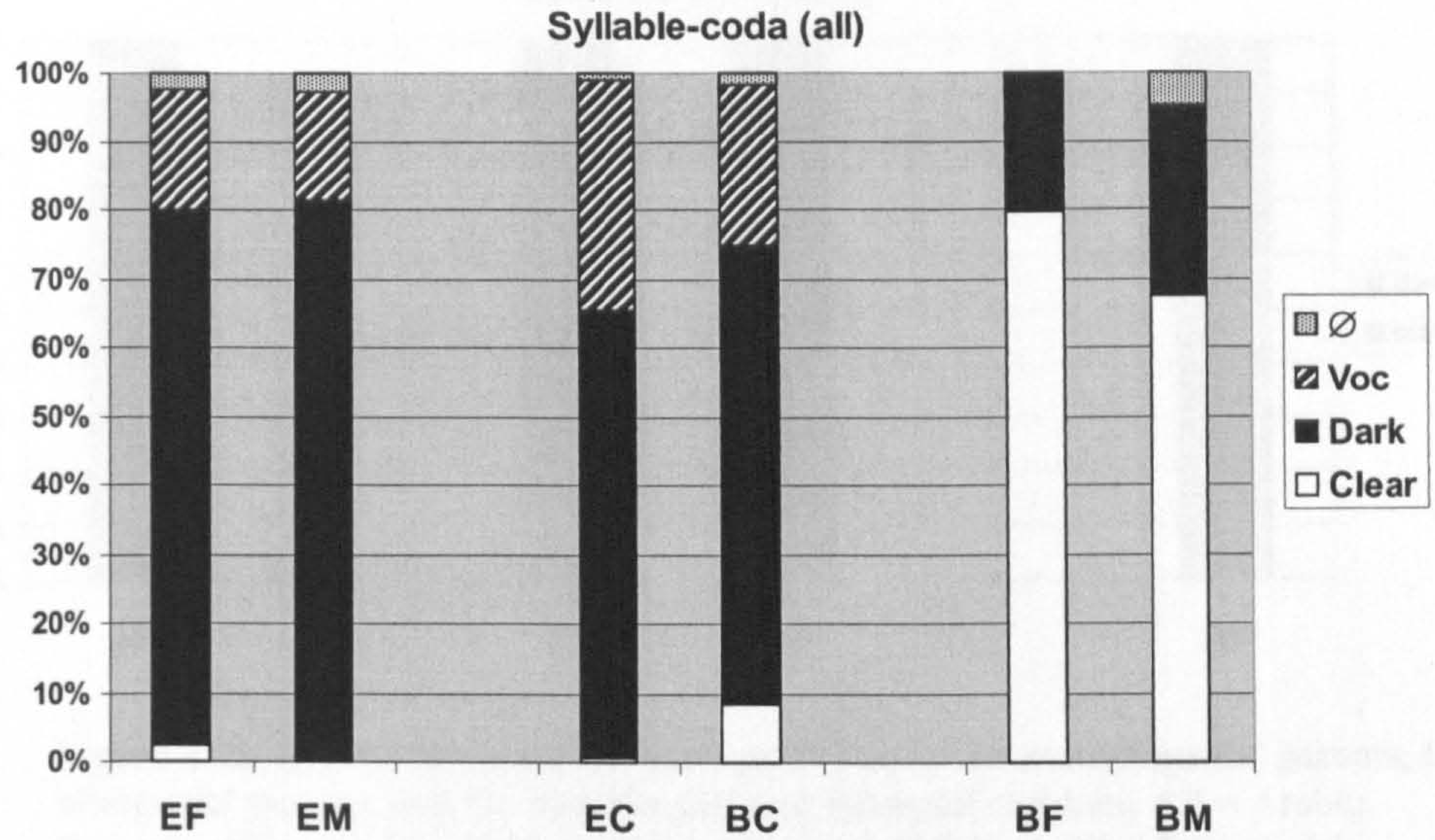


Figure 3.21: Results for syllable-coda and syllabic /l/ in English by the monolinguals' parents, the bilinguals' parents, and the monolingual English and bilingual children. EF = English Females; EM = English Males; EC = English children; BC = bilingual children; BF = Bilinguals' parents (Females); BM = Bilinguals' parents (Males). N = 1065.

3.8 Arabic results

3.8.1 Syllable-onset and coda

3.8.1.1 auditory analysis

The results for this part of the analysis can be described very briefly, since analysis showed a categorical production of clear /l/'s in both syllable-onset and coda positions for all the bilinguals' and the monolinguals' parents (Figure 3.22 and Table 3.13).

Such results constitute a strong piece of evidence that the bilingual subjects have acquired the correct patterns for Arabic /l/ production and that their productions do not show any signs of interference from English, since there are no dark or vocalised /l/'s in final position. Both adults and children occasionally omitted word-final /l/'s in running speech (Table 3.13). There were no obvious differences between the three age groups with respect onset and coda /l/'s, and, apart from omissions, other realisations by the children included [ɾ] and [m] substitutions, e.g. [faɾat] for [falat] 'he let go' [mˈbuːme] for [lˈbuːme] 'the owl'.

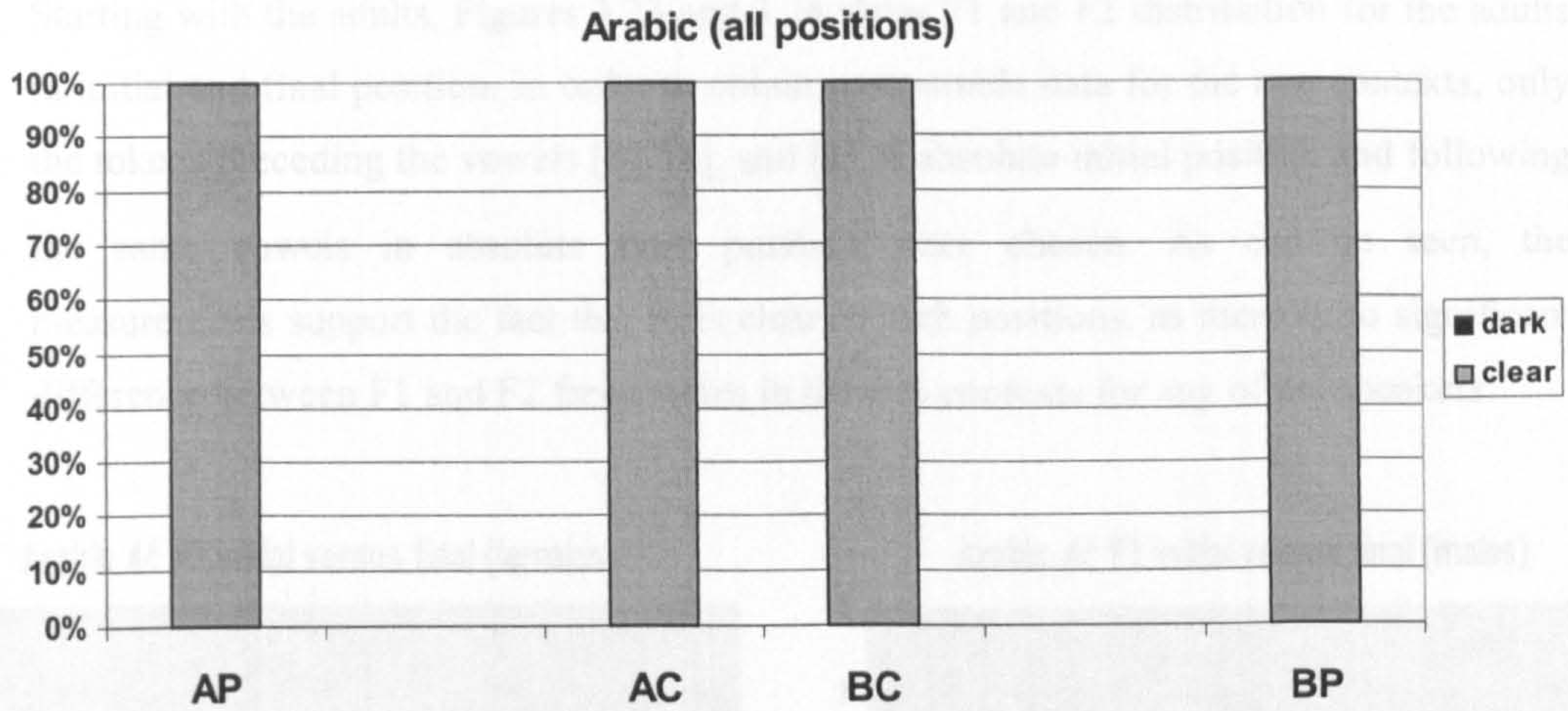


Figure 3.22: Results for Arabic /l/ in all positions by the monolinguals' parents, the bilinguals' parents, and the monolingual and bilingual children. AP = Arabic Parents; AC = Arabic Children; BC = Bilingual Children; BP = Bilinguals' Parents. N = 1284.

Table 3.13: Detailed results for Arabic /l/ patterns produced by all 14 subjects in onset (O) and coda (C) positions.

	AF5			AF7			AM5			AM10		
	O	**C**	**N**	**O**	**C**	**N**	**O**	**C**	**N**	**O**	**C**	**N**
Clear	69	43	112	50	41	91	55	54	109	59	62	121
Dark	0	0	0	0	0	0	0	0	0	0	0	0
Ø	0	0	0	0	1	1	0	0	0	0	2	2
Total	**69**	**43**	**112**	**50**	**42**	**92**	**55**	**54**	**109**	**59**	**64**	**123**

	BF5			BF7			BM5			BM7		
	O	**C**	**N**	**O**	**C**	**N**	**O**	**C**	**N**	**O**	**C**	**N**
Clear	40	37	77	46	46	92	52	39	91	50	46	96
Dark	0	0	0	0	0	0	0	0	0	0	0	0
Ø	0	0	0	0	0	0	0	0	0	0	0	0
Total	**40**	**37**	**77**	**46**	**46**	**92**	**52**	**39**	**91**	**50**	**46**	**96**

	A5			A7			A10			B5			B7			B10		
	O	**C**	**N**	**O**	**C**	**N**	**O**	**C**	**N**	**O**	**C**	**N**	**O**	**C**	**N**	**O**	**C**	**N**
Clear	36	30	66	61	38	99	54	46	100	23	16	39	40	44	84	59	43	102
Dark	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Ø	0	0	0	0	0	0	0	1	1	0	0	0	0	3	3	0	1	1
other	0	0	0	1	0	1	1	0	1	0	0	0	1	4	5	0	0	0
Total	**36**	**30**	**66**	**62**	**38**	**100**	**55**	**47**	**102**	**23**	**16**	**39**	**40**	**51**	**91**	**59**	**44**	**103**

3.8.1.2 acoustic analysis: Adults

Starting with the adults, Figures 3.23 and 3.24 show F1 and F2 distribution for the adults in initial and final position. In order to obtain comparable data for the two contexts, only the tokens preceding the vowels [i:], [e], and [a] in absolute initial position and following the same vowels in absolute final position were chosen. As can be seen, the measurements support the fact that /l/ is clear in both positions, as there is no significant difference between F1 and F2 frequencies in the two contexts for any of the speakers.

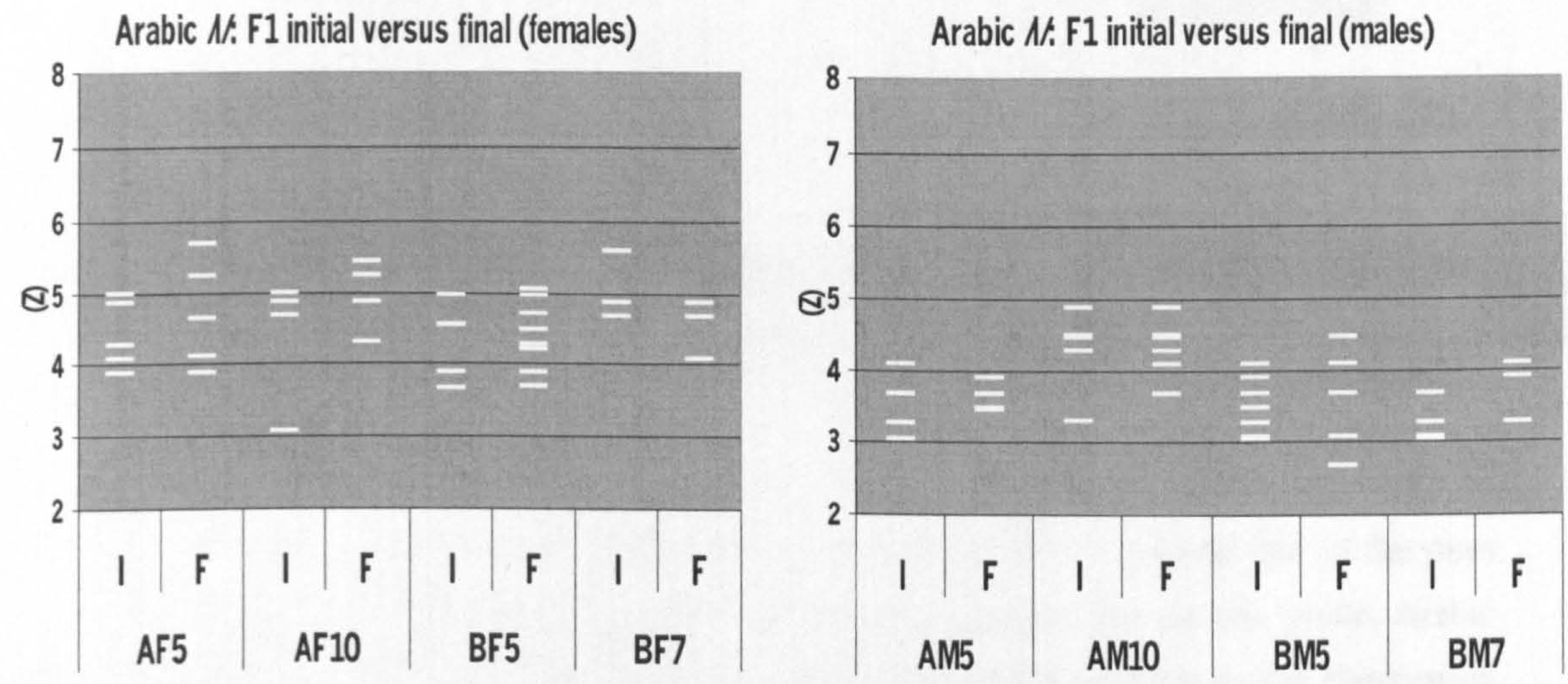


Figure 3.23: Bark-scaled F1 measurements for syllable-initial and syllable-final /l/ in Arabic produced by the monolinguals' parents and the bilinguals' parents. White dashes indicate perceptually clear [l]'s. N = 86.

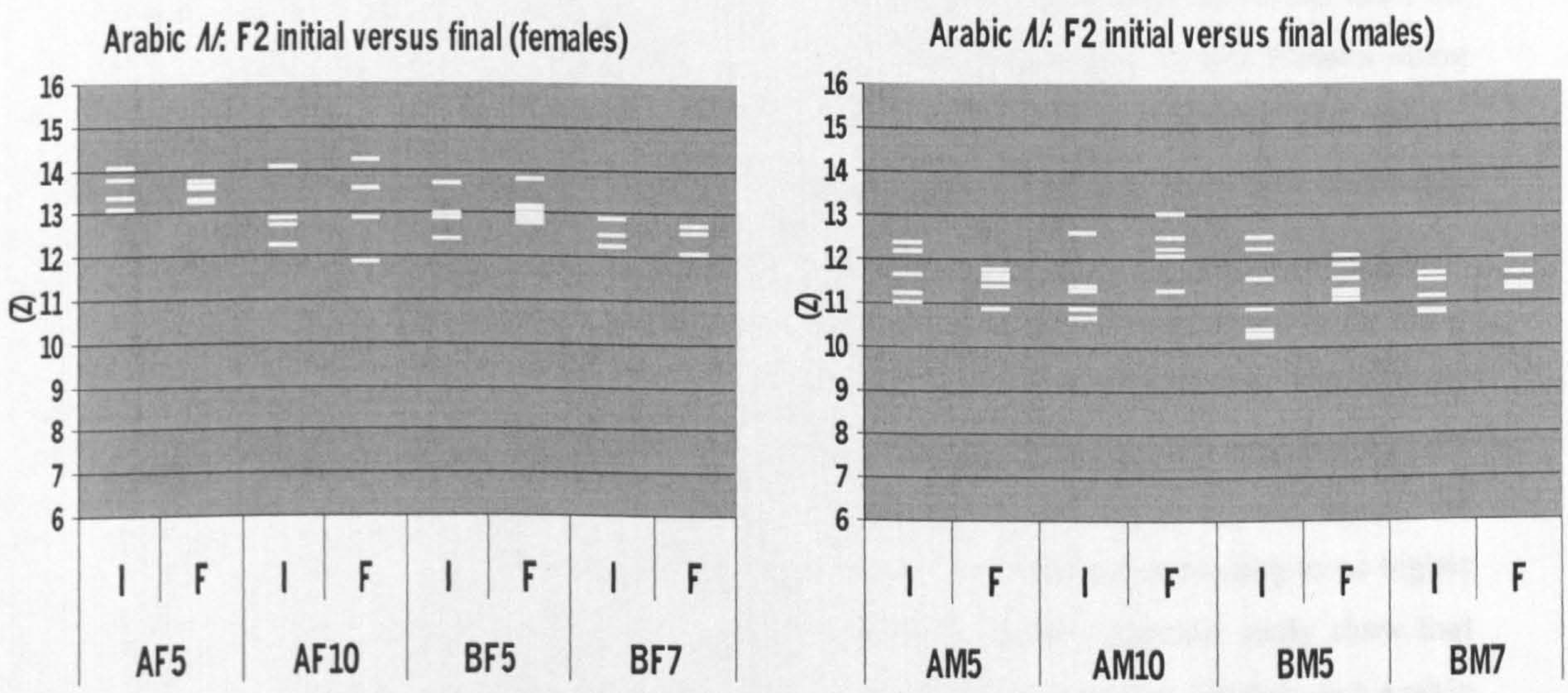


Figure 3.24: Bark-scaled F2 measurements for syllable-initial and syllable-final /l/ in Arabic produced by the monolinguals' parents and the bilinguals' parents. White dashes indicate perceptually clear [l]'s. N = 81.

Interestingly, final clear [l]'s did not necessarily have lower F2 frequencies than initial [l]'s in comparable vocalic contexts, despite the general tendency for /l/'s in coda position to be darker than in onsets (Sproat & Fujimura, 1993). This point is illustrated in Figure 3.25, which shows a spectrogram of the words [li:fe] 'sponge' and [fi:l] 'elephant' by BM10, whereby final F2 is higher than initial F2. F2 for [l] in [li:fe] is 11.63Z, while that for [fi:l] is 11.99Z. 1

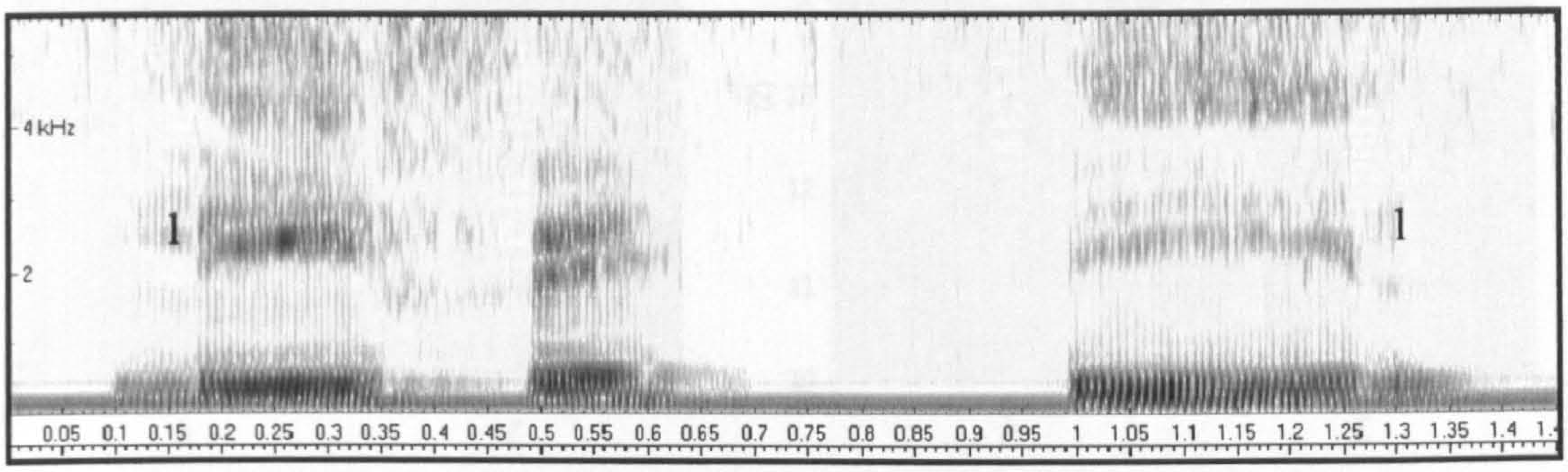


Figure 3.25: Spectrogram of the words [li:fe] 'sponge' (left) and [fi:l] 'elephant' (right) as produced by BM10.

My own perception of English and Arabic clear [l]'s in general and of the ones produced by the subjects in this study specifically suggests that on the whole, Arabic clear [l]'s sound clearer than English clear [l]'s. In order to investigate this observation using data from this study, the only context I could examine was initial position for females, since the monolingual English males produce a considerable number of dark [ɫ] tokens in this position. A subset of the data consisting of initial clear [l] tokens from the monolingual English females on the one hand, and the monolingual Arabic females along with the bilinguals' mothers[5] on the other was extracted for comparison. The vocalic context included the vowels [i:], [e], and [a] in order to control data from both languages. Figure 3.26 shows F1 and F2 distributions for clear initial [l]'s in Arabic and English. While there was no significant difference between the F1 distributions for clear [l] in the two languages ($p = 0.2$), the F2 range for initial clear [l]'s in Arabic was significantly higher than that of English clear [l]'s despite the expected overlap (t-test significant at $p < 0.01$). While F2 frequencies in Arabic were concentrated in the 12 to 14Z range, the English ones had a wider range starting from as low as 10.84Z and extending to no higher than 13.35Z. Although the number of tokens is small, results from this study show that there might be finer differences within the 'clear' category between English and Arabic

---

4 The decision to include the bilinguals' mothers was made after checking that there was no significant difference between their F2 values for initial /l/ and those of the monolingual females.

and that these should be taken into consideration when describing the quality of /l/in the two languages.

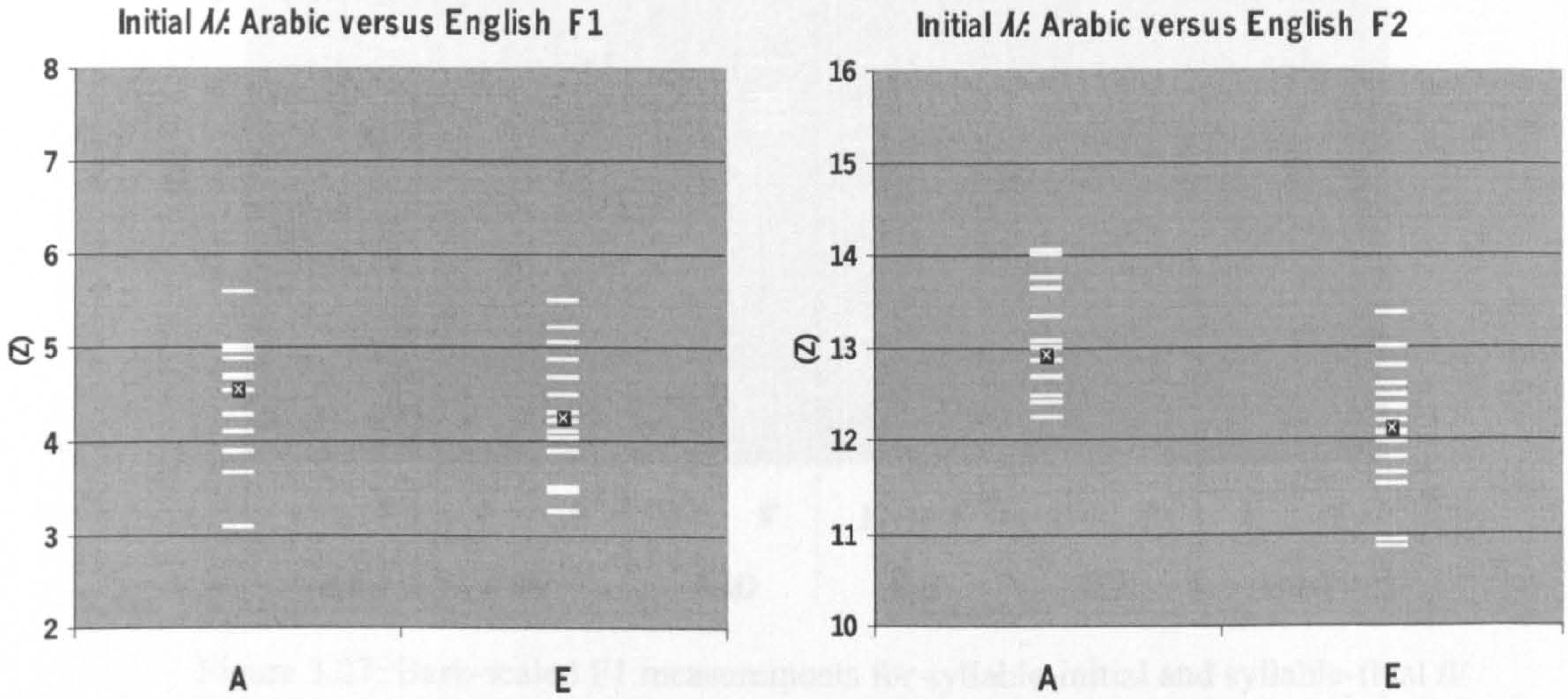


Figure 3.26: Bark-scaled F1 and F2 measurements for syllable-initial clear [l] in Arabic and English produced by the monolingual Arabic mothers and the bilinguals' mothers (A), and the monolingual English mothers (E). Crosses indicate means. N = 69.

3.8.1.3 acoustic analysis: children

Moving on to the acoustic results of the children, Figures 3.27 and 3.28 show F1 and F2 distribution for the children in initial and final position. In order to obtain comparable data for the two contexts, only the tokens preceding the vowels [i:], [e], and [a] in initial position and following the same vowels in final position were chosen. F1 and F2 frequencies were much more variable for the children than for the adults (Figures 3.23 and 3.24), and two of the bilingual children (B7 and B10) do seem to have lower F1 and F2 frequency ranges than their monolingual counterparts, but there were not enough tokens to test the significance of the difference between the two groups. It nevertheless seems, although both bilingual and monolingual groups are producing clear [l]'s in initial and final position, that there are might be more subtle differences in the quality of the 'clear' [l]'s produced, especially with respect to F2 in final position for B7 and B10 (Figure 3.28).

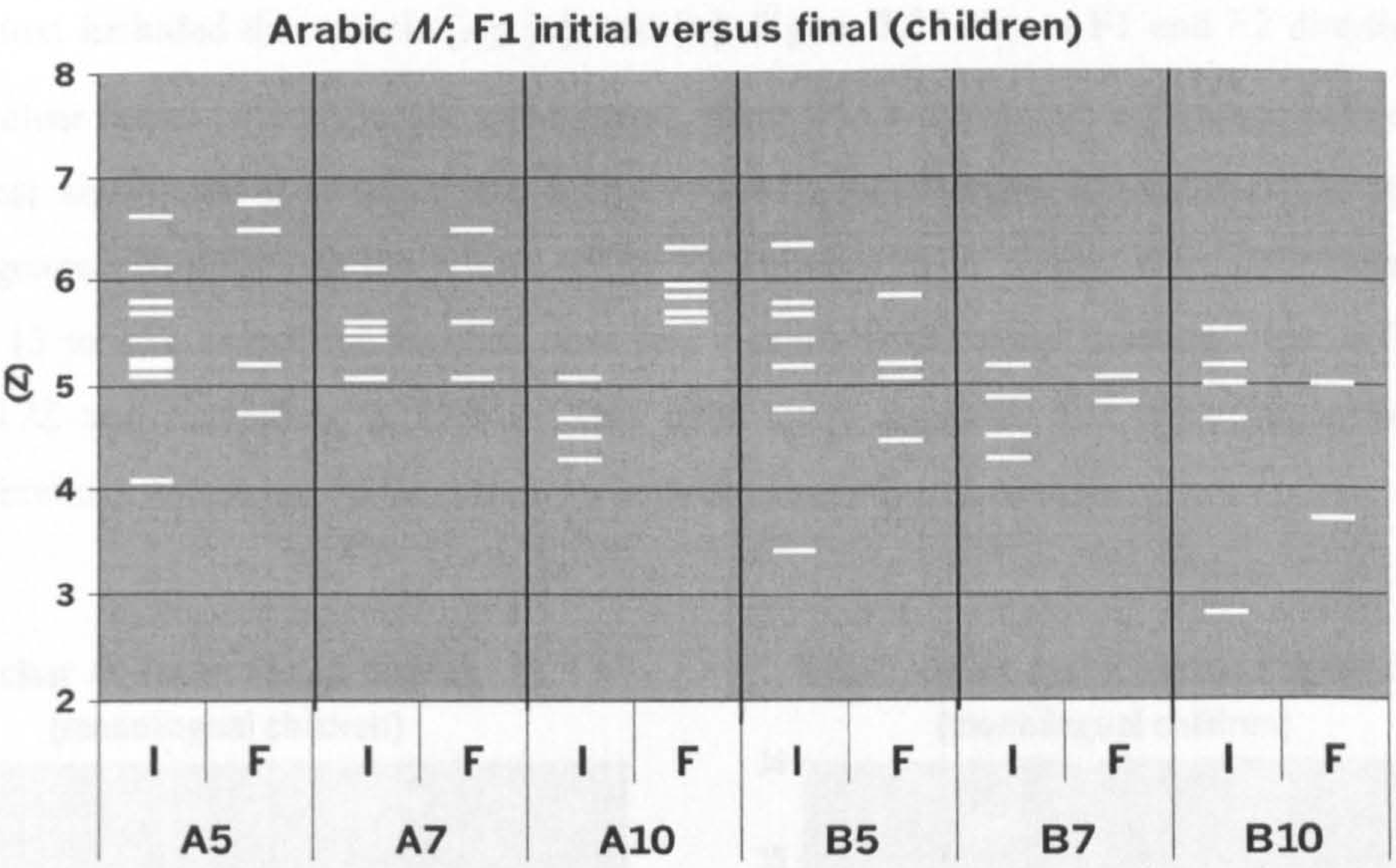


Figure 3.27: Bark-scaled F1 measurements for syllable-initial and syllable-final /l/ in Arabic produced by the monolingual and bilingual children. White dashes indicate perceptually clear [l]'s. N = 63.

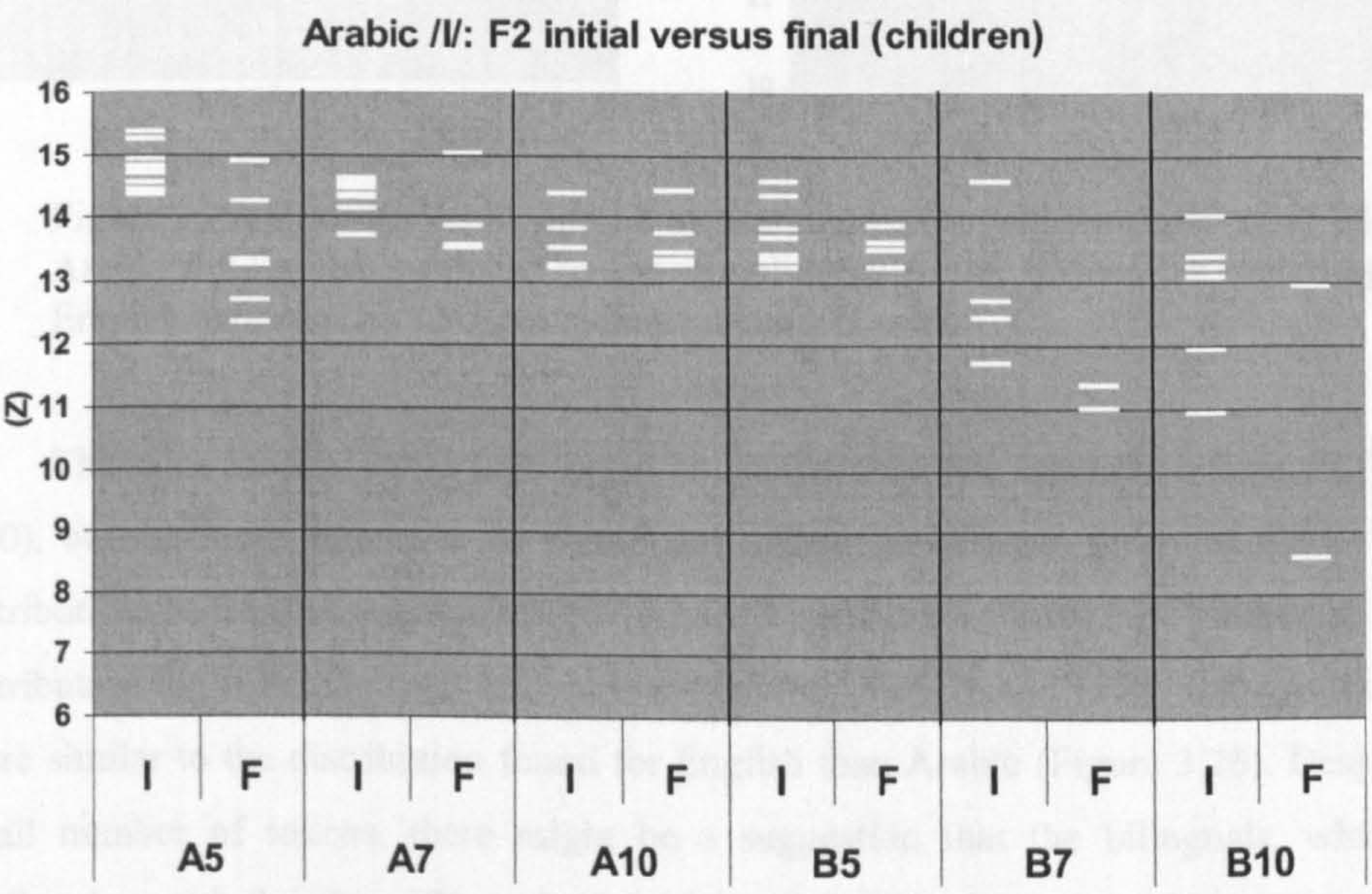


Figure 3.28: Bark-scaled F2 measurements for syllable-initial and syllable-final /l/ in Arabic produced by the monolingual and bilingual children. White dashes indicate perceptually clear [l]'s. N = 61.

In order to test the tendency for clear [l]'s in Arabic to be somewhat 'clearer' than English clear [l] which was found for the adults (Figure 3.26), a similar examination was carried out for the children, this time using data only from the monolingual Arabic and monolingual English children while analysing the bilinguals' data separately. The vocalic context included the vowels [i:], [e], and [a]. Figure 3.29 shows F1 and F2 distributions for clear initial [l]'s in Arabic and English. there was a significant difference between F1 (t-test significant p < 0.01) and F2 (p = 0.01) distributions for clear [l] in the two languages. Similarly to the adults, while F2 frequencies in Arabic were concentrated in the 13 to 15Z range, the English ones had a much wider range starting from as low as 10.17Z and extending to 15.74Z. This once again suggests that there might be finer differences within the 'clear' category between English and Arabic.

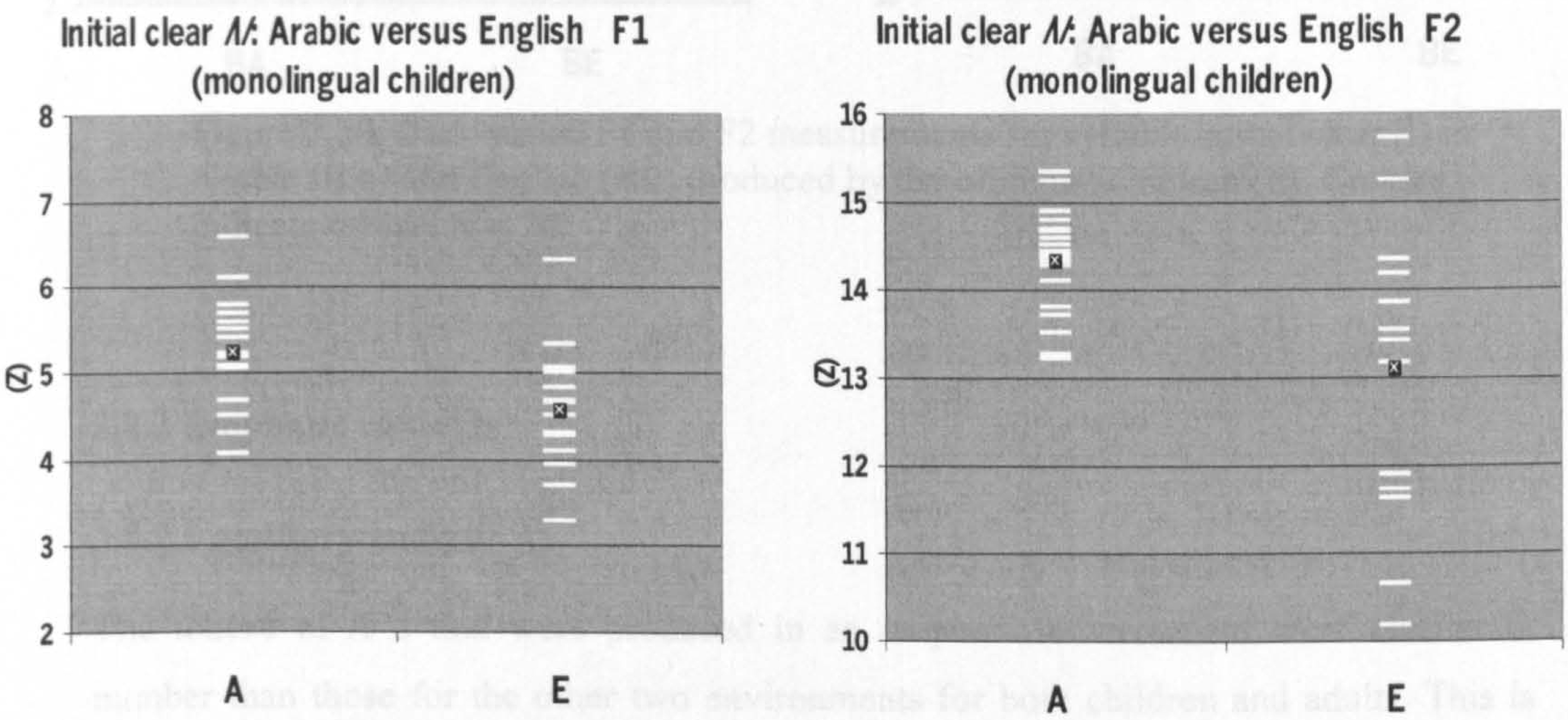


Figure 3.29: Bark-scaled F1 and F2 measurements for syllable-initial clear [l] in Arabic and English produced by the monolingual Arabic (A)and the monolingual English children (E). Crosses indicate means. N = 69.

The same tendency was then tested on comparable data from the bilinguals (Figure 3.30), but this time there was no significant difference between either of the F1 or F2 distributions in English and Arabic (p = 0.1 for F1 and p = 0.08 for F2). Moreover, the F2 distribution for both languages had wide ranges (between 11 and 15Z) and was therefore more similar to the distribution found for English than Arabic (Figure 3.26). Despite the small number of tokens, there might be a suggestion that the bilinguals, while still conforming with the clear [l]'s in both initial and final position in Arabic, are producing formant frequencies that are closer to English clear [l]'s than Arabic ones; this in turn suggests that they might be using different articulatory strategies for their Arabic clear [l]'s than the monolingual Arabic children. Further investigation using more data is needed to support this claim.

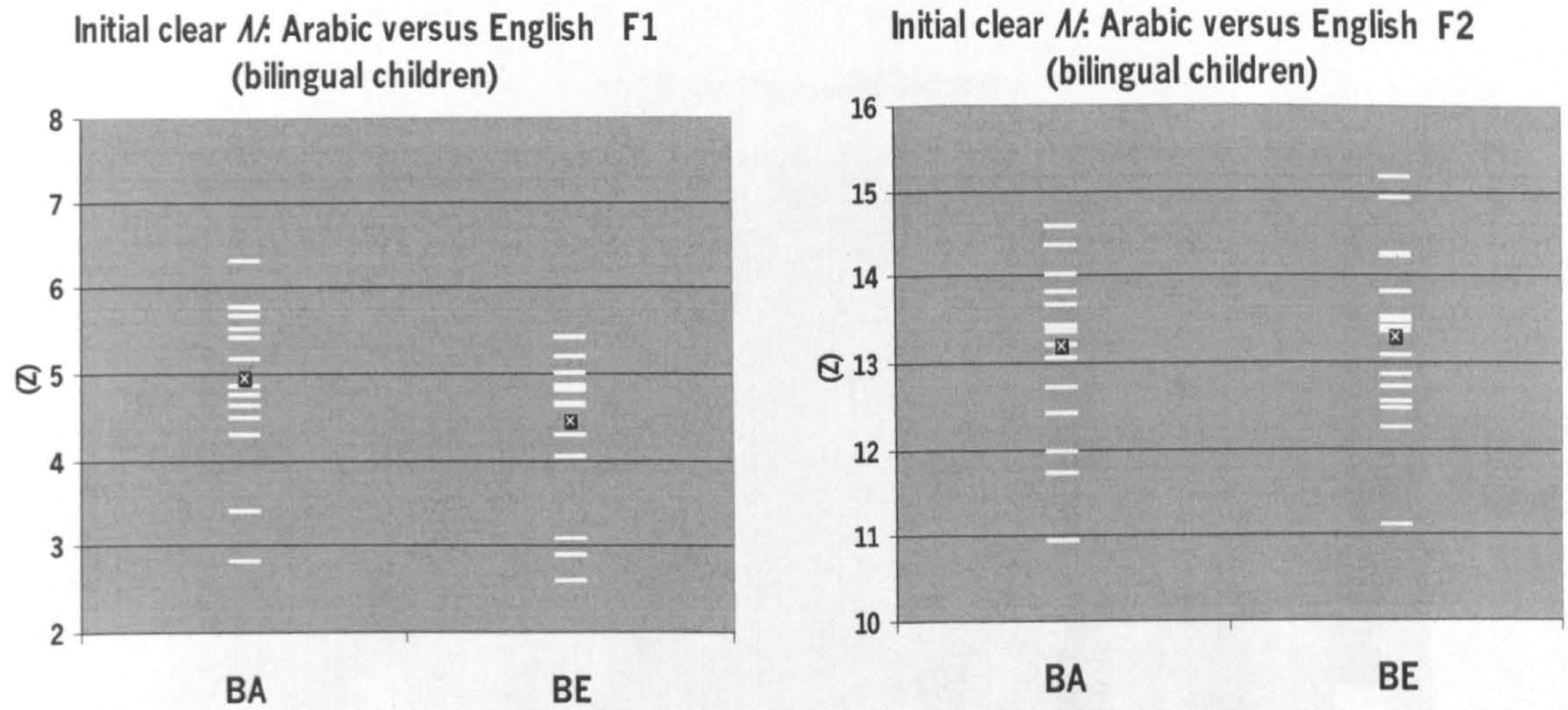


Figure 3.30: Bark-scaled F1 and F2 measurements for syllable-initial clear [l] in Arabic (BA) and English (BE) produced by the bilingual children (E). Crosses indicate means. N = 34.

3.8.2 Emphatic contexts

3.8.2.1 auditory analysis

The tokens of /l/'s that were produced in an emphatic environment were smaller in number than those for the other two environments for both children and adults. This is due to the fact that words with such a combination (an emphatic sound + /l/ in the same word) are not frequent in the language, and even the existent ones do not always involve a spread of emphasis from the emphatic consonant to the neighbouring /l/ sound. Factors such a directionality (leftwards or rightwards) and degree of spread vary across dialects and speakers within dialects (see Section 3.2.1). Apart from that, it is more difficult to find emphatic tokens for children since, as mentioned in Section 3.2.4, emphasis is one of the sound features that are acquired very late in children and often remain only partially developed until the age of 14. For this reason, only 310 tokens were found for all 14 speakers (three monolingual Arabic children, three bilingual children, and eight adults). Figures 3.31 and 3.32 and Table 3.14 show the patterns for /l/ production in emphatic environments by both adults and children.

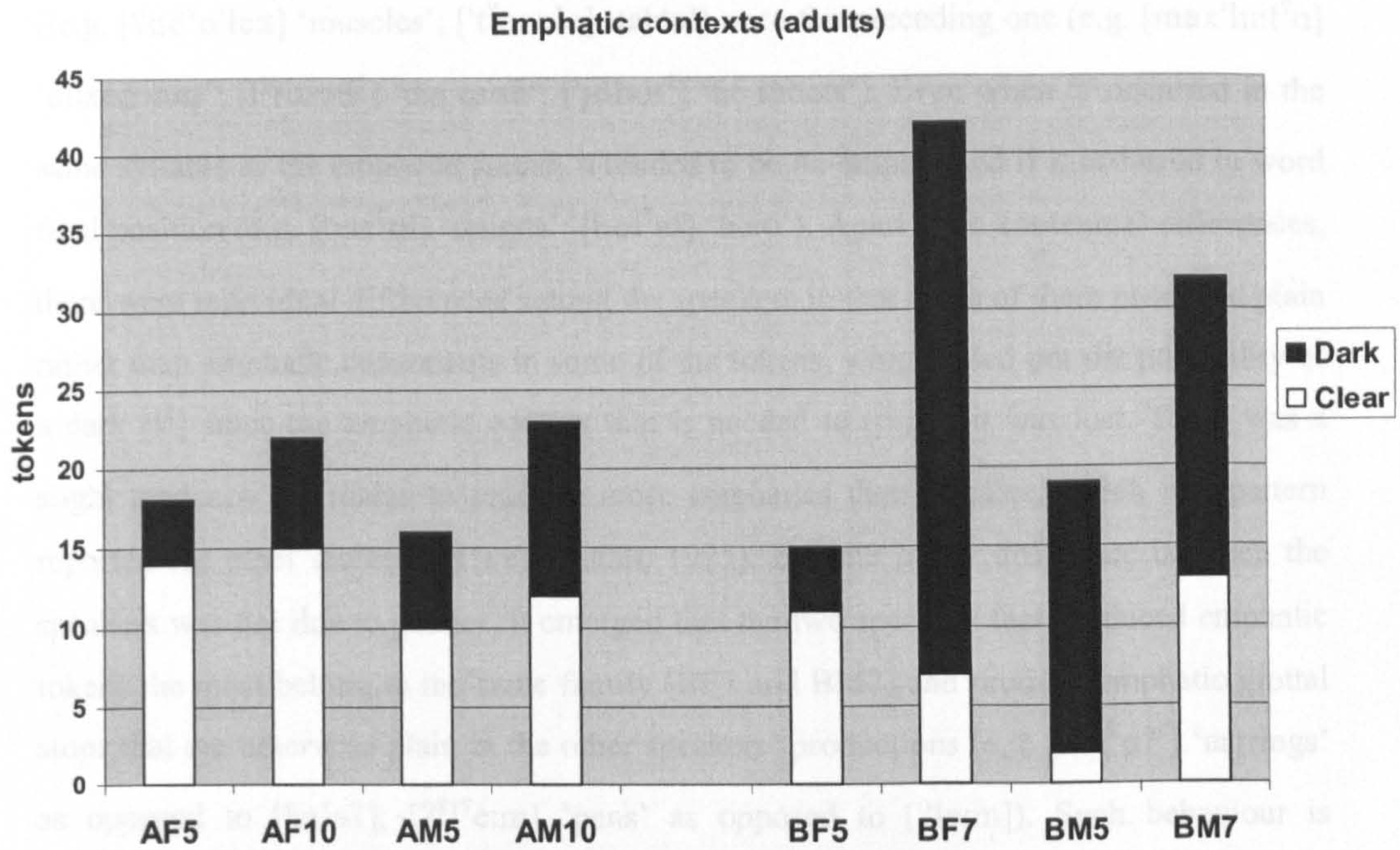


Figure 3.31: Results for /l/ in emphatic contexts in Arabic by the monolinguals' parents and the bilinguals' parents. N (tokens) = 187.

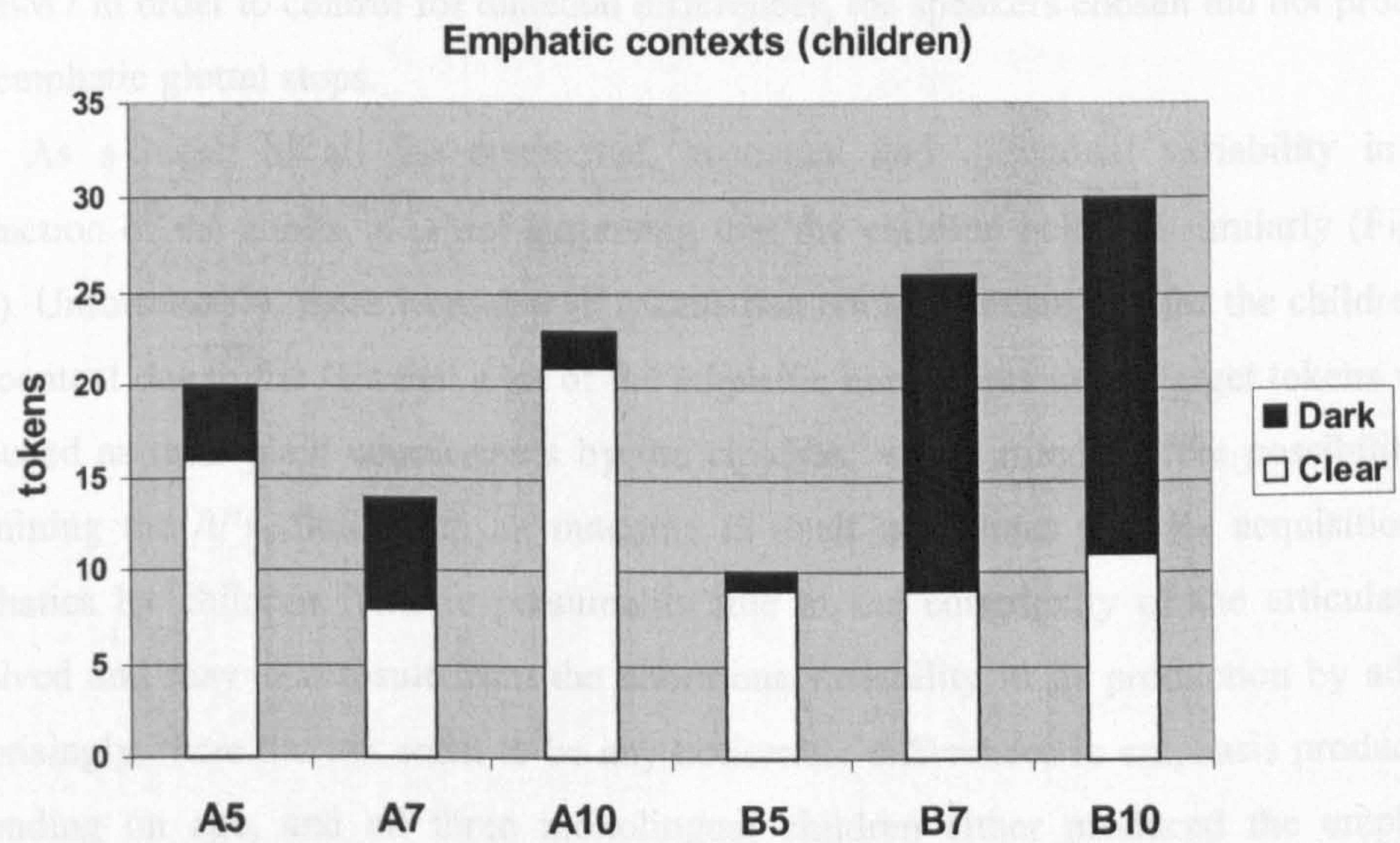


Figure 3.32: Results for /l/ in emphatic contexts in Arabic by the monolingual Arabic and the bilingual children. N = 123.

Starting with the adults, it is interesting to note that none of the speakers produces categorical dark [lˤ]'s in all the target tokens. There are various reasons for this finding. First, not all the emphatic contexts in the target words triggered emphasis spread to the

/l/'s in those tokens. For instance, emphasis did not always spread to the next syllable ([e.g. [ʕadˤaˈleːt] 'muscles'; [ˈtˤaːwle] 'table') or to the preceding one (e.g. [maxˈluːtˤa] 'mixed nuts'; [lˈʔarədˤ] 'the earth'; [ˈjɪlbutˤ] 'he shoots'). Even when /l/ occurred in the same syllable as the emphatic sound, it tended to be de-emphasised if it occurred in word final position (e.g. [basˤal] 'onions'; [batˤɐl] 'hero'). Apart from contextual differences, there were individual differences among the speakers in that some of them produced plain rather than emphatic consonants in some of the tokens, which ruled out the possibility of a dark [lˤ] since the emphatic context that is needed to trigger it was lost. There was a slight tendency for males to produce more emphatics than females, which is a pattern reported for other dialects as well (Kahn, 1975). But the major difference between the speakers was not due to gender. It emerged that the two speakers that produced emphatic tokens the most belong to the same family (BF7 and BM7) and produce emphatic glottal stops that are otherwise plain in the other speakers' productions (e.g. [ħalˤaʔˤ] 'earrings' as opposed to [ħalaʔ]; [ʔˤlˤeːm] 'pens' as opposed to [ʔleːm]). Such behaviour is particular to the accent of certain localities within Beirut, where the historic uvular plosive [q] that changed into [ʔ] in the Lebanese dialect is still produced with emphasis, therefore [ʔˤ]. Although the Lebanese families were chosen from the same locality as BF7 and BM7 in order to control for dialectal differences, the speakers chosen did not produce any emphatic glottal stops.

As a result of all the contextual, accentual and individual variability in the production of the adults, it is not surprising that the children behaved similarly (Figure 3.32). Unfortunately, there were few /l/ tokens that could be examined for the children in this context due to the fact that a lot of the emphatic consonants in the target tokens were produced as their plain counterparts by the children, which ruled out the possibility of examining the /l/'s. Still, such an outcome in itself underlines the late acquisition of emphatics by children. This is presumably due to the complexity of the articulations involved and may also result from the enormous variability in its production by adults. Surprisingly, there did not seem to be any noticeable differences in emphasis production depending on age, and all three monolingual children either produced the emphatic consonants in the target tokens but kept the /l/'s plain (e.g. [basˤɐl] 'onions'; [ʕadˤɐˈleːt] 'muscles'), or produced plain sounds throughout (e.g. [dalːit] for [dˤalˤːit] 'she stayed'). As for the bilinguals, while B5 behaves more or less similarly to A5, B7 and B10 produce more emphatic tokens. The most likely reason for this is the fact that they are BF7 and BM7's children and therefore produced emphatic glottal stops as well (e.g. [ˈʔˤalˤʕa] 'fortress'; [ˈlˤaʔˤɪt] 'she found'; [ˈʔˤaːlɪt] 'she found'). B7 and B10 have therefore acquired emphasis as present in the input that they receive, along with the accent-specific feature of their parents.

Table 3.14: Detailed results for Arabic /l/ patterns produced by all 14 subjects in emphatic contexts.

	AF5			AF7			AM5			AM10		
	read	story	N	read	story	N	read	story	N	read	story	N
Clear	4	10	14	5	10	15	4	6	10	4	8	12
Med	0	1	1	0	2	2	0	3	3	0	4	4
Dark	3	0	3	2	3	5	3	0	3	3	4	7
Total	7	11	18	7	15	22	7	9	16	7	16	23
	BF5			BF7			BM5			BM7		
	read	story	N	read	story	N	read	story	N	read	story	N
Clear	5	6	11	2	5	7	1	1	2	6	7	13
Med	0	2	2	0	12	12	2	4	6	5	4	9
Dark	1	1	2	9	14	23	3	8	11	2	8	10
Total	6	9	15	11	31	42	6	13	19	13	19	32

	A5			A7			A10		
	pic	story	N	pic	story	N	pic	story	N
Clear	4	12	16	2	6	8	4	17	21
Med	0	2	2	0	2	2	0	1	1
Dark	0	2	2	2	2	4	0	1	1
Total	4	16	20	4	10	14	4	19	23

	B5			B7			B10		
	pic	story	N	pic	story	N	pic	story	N
Clear	7	2	9	5	4	9	9	2	11
Med	0	0	0	0	1	1	0	8	8
Dark	0	1	1	8	8	16	5	6	11
Total	7	3	10	13	13	26	14	16	30

3.8.2.2 acoustic analysis: adults

Figures 3.33 and 3.34 show F1 and F2 distribution for the adults' production of /l/ tokens in emphatic contexts for words produced in isolation. Due to the small number of tokens, the preceding and/or following vocalic contexts were not controlled and the main aim will be to show the general patterns for the /l/'s that were categorised as clear as opposed to the ones that were categorised as dark in this context. No discernable pattern can be detected for F1 differences between clear (white dashes) and emphatic (dark dashes) tokens apart from a general tendency for emphatic [lˤ] in the male data to have a higher F1 (Figure 3.29). However, F2 remained a strong predictor of darkness for all speakers as can be shown in Figure 3.30. F2 frequencies are considerably lower in emphatic [lˤ]'s than in clear [l]'s. While the majority of clear [l]'s in this context had similar frequencies to the ones found in non-emphatic contexts, some clear [l]'s had frequencies that were lower than any of the ones found in plain contexts in Figure 3.24 (e.g. AM5 and AM10). This may suggest that some of the /l/'s that are heard as 'clear' in an emphatic context might still be darker than clear [l]'s in plain environment, as there is a continuum of clearness and darkness involved.

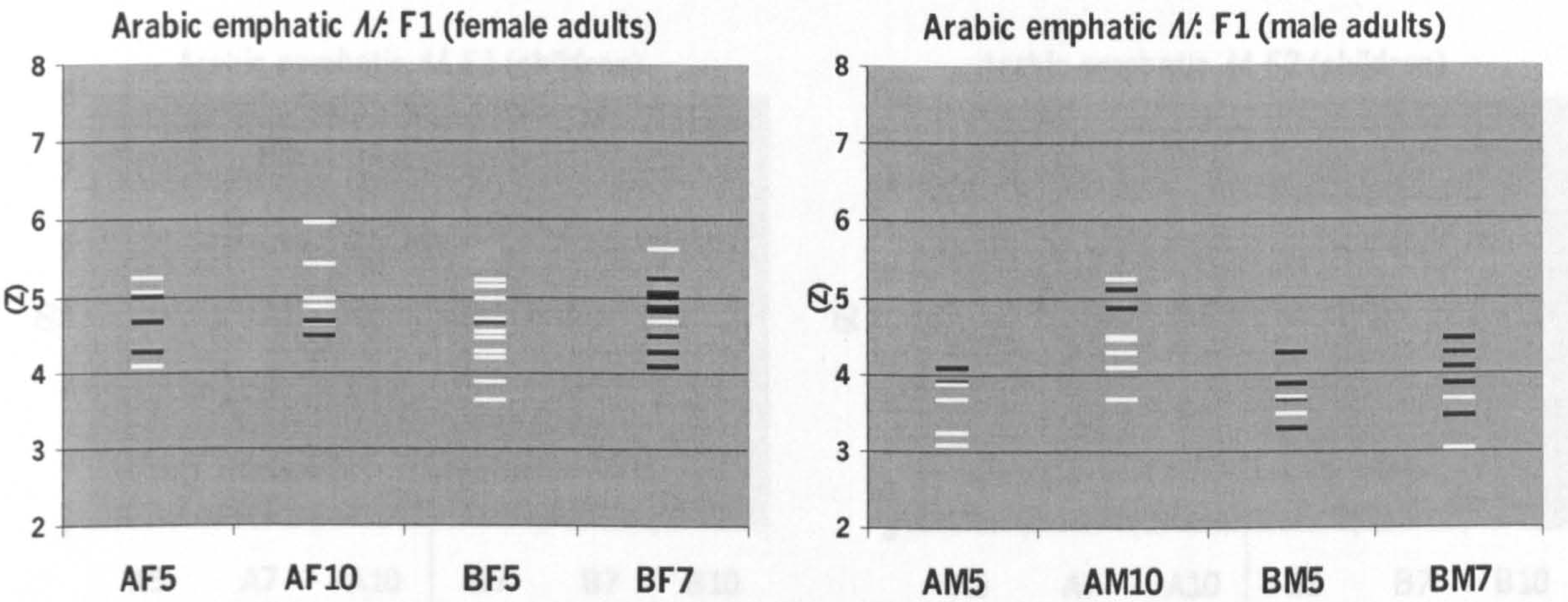


Figure 3.33: Bark-scaled F1 measurements for /l/ in emphatic context in Arabic produced by the females (left) and the males (right). White dashes indicate perceptually clear [l]'s, black dashes indicate perceptually dark [lˤ]'s.

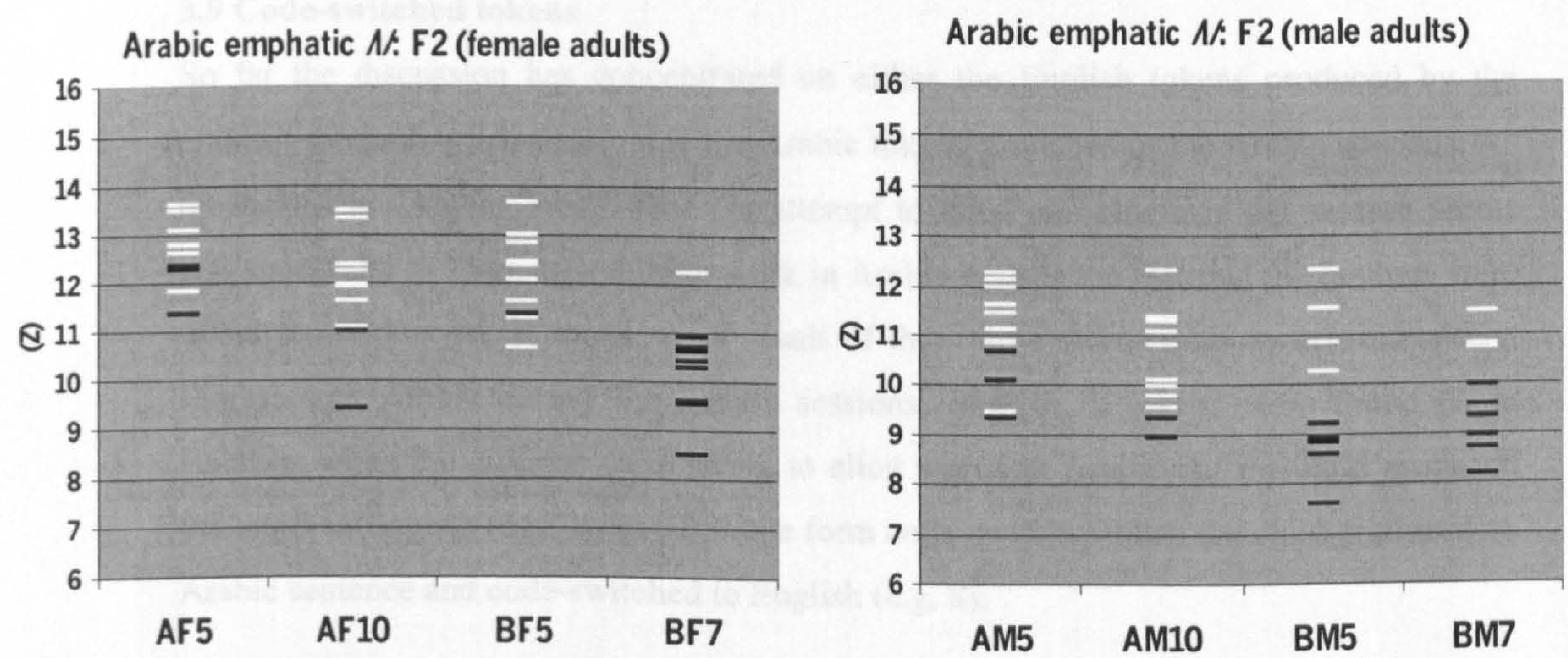


Figure 3.34: Bark-scaled F2 measurements for /l/ in emphatic contexts in Arabic produced by the females (left) and the males (right). White dashes indicate perceptually clear [l]'s, black dashes indicate perceptually dark [lˤ]'s.

3.8.2.3 acoustic analysis: children

Figure 3.35 shows the F1 and F2 distribution for the children's production of /l/ tokens in emphatic contexts. As mentioned for the adults, it is mainly F2 that shows a discernible pattern in terms of the clear/dark distinction. F2 frequencies for emphatic [lˤ]'s are considerably lower than those for plain [l]'s. However, when comparing B7's F1 measurements in emphatic contexts with those obtained for his clear [l] production in Figure 3.23, one can see that F1 tends to be higher in emphatic than in plain contexts (though the difference was not significant).

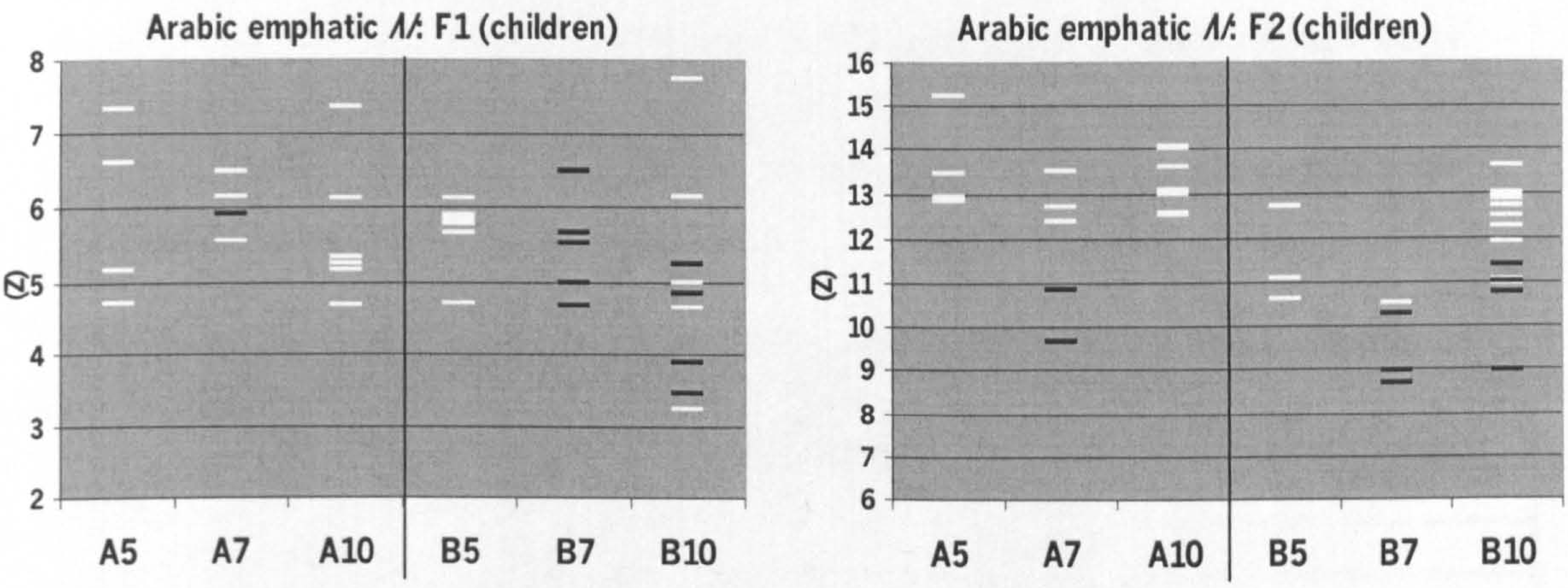


Figure 3.35: Bark-scaled F1 and F2 measurements for /l/ in emphatic contexts in Arabic produced by the monolingual and bilingual children. White dashes indicate perceptually clear [l]'s, black dashes indicate perceptually dark [lˤ]'s.

3.9 Code-switched tokens

So far the discussion has concentrated on either the English tokens produced by the children in the English sessions or the Arabic tokens produced in the Arabic sessions. As mentioned in Chapter Two, while the attempt to elicit one language per session seems was successful in English, it did not work in Arabic despite the fact that the mothers were asked to conduct the sessions. As a result of that, the children code-switched between English and Arabic during the Arabic sessions. English /l/ tokens were found (i) in isolation when the mothers were trying to elicit words in Arabic but the child produced the target in English (e.g. 7), or (ii) in the form code-switches when the child produced an Arabic sentence and code-switched to English (e.g. 8).

(7) Mother (pointing at a kettle): [ʃu haɪda]?
What that (masc.)?
'What is that?'
Child: [ˈkɛtəl]
<u>KETTLE</u>

(8) Child (describing a an action): [nɑtˤtˤo bɪl puːl]
jump-past-3^{rd} pers.pl. in <u>POOL</u>
'they jumped in the pool'

Unfortunately, the /l/ tokens were not as numerous but still needed to be examined separately from the data presented above due to the difference in patterns that the children exhibited in these tokens. While in the English sessions the three subjects showed evidence of having acquired dark or vocalised word-final /l/'s, the majority of English tokens that these subjects produced in the Arabic sessions had clear [l]'s in all word-positions, which is the pattern that is expected for Arabic and that was produced by the bilinguals' parents in English. It is interesting to note that most of the lexical items are the same ones that the bilinguals produced in the English sessions, which allowed direct comparison of two pronunciations of the 'same' word. Table 3.15 shows transcripts of the English tokens produced by the children in the Arabic sessions and, where possible, in the English sessions.

Table 3.15: English target words produced by the bilinguals during the Arabic sessions (left) and during the English sessions (right)

	Gloss	Produced in Arabic sessions	Produced in English sessions
B5	pool	pʰʉ̯əl	pʰʉʊ
	castle	ˈkʰa̯sədə	ˈkʰasɫ̩
	marbles	ˈma̯rbəz̥	ˈmɑːbɫ̩z̥
	kettle	ˈkɛtəl	ˈkʰɛtəɫ
	teletubbies	ˈtejtʌbiː	ˈtʰɛlɪtʌbiː
	elbow	ˈɛlbo	ˈɛʊbə
	bottle	ˈbɔtəl	ˈbɔtəɫ
	all	ɔːl	-

	Gloss	Produced in Arabic sessions	Produced in English sessions
B7	purple	ˈpɜːpəl	ˈpɜːpɵ
	muscle	ˈmʌsˤl̩	ˈmʌsɫ̩z̥
	football	fətˈboːl	ˈfuʔbəʊ
	bottle	ˈbɔtəl	ˈbɒtɐ
	elbow	ˈɛlbɔ	ˈɛɫbɔᵊ
	nails	neɪəɫz̥	neɪɫz̥
	while	waɪəɫ	-
	little	ˈlɪtəl	-
	called	kɔːld	-

	Gloss	Produced in Arabic sessions	Produced in English sessions
B10	castle	ˈkasl̩	ˈkʰaso
	pool	puːˀl	pʰʉɤ
	football	fotˈboːl	fʊʔbɔʊ
	baseball	ˈbe˞sbəl	-
	couple	ˈkʌpəl	-

The transcriptions in Table 3.8 reveal that the subjects did not only change their pronunciations of /l/'s when producing the 'same' tokens depending on the language session, but also changed other sound features that are normally produced differently in the two languages. For instance, words with initial voiceless stops often had much shorter VOT when produced in the Arabic sessions than in the English ones (e.g. 'kettle' and 'teletubbies' for B5, 'castle' and 'pool' for B10). Such behaviour can be explained by the fact that voiceless stops in Arabic are normally unaspirated (Chapter Five), while their English counterparts tend to be aspirated. Another noticeable change is in the vowels produced in the tokens from the Arabic sessions. Open front vowels were sometimes raised, fronted, or backed so that they conveyed a quality that was closer to how they would normally be produced in Arabic (e.g. [ˈkʰa̹sədə] 'castle' and [ˈma̹ɾbəz̧] 'marbles' by B5; [fətˈboːl] 'football' by B7 and B10, and [puːl] 'pool' by B10). [ˈma̹ɾbəz̧] also shows the emergence of a post-vocalic /r/ that is otherwise lacking in the bilingual subjects' productions in English (Chapter Four). Yet another pattern is the insertion of a schwa in tokens with otherwise syllabic /l/'s regardless of the /l/ realisation ([kʰa̹sədə] 'castle'; [ˈkɛtəl] 'kettle'; and [ˈbɔtəl] 'bottle' by B5; [ˈpɜːpəl] 'purple'; [ˈbɔtəl] 'bottle'; and [ˈlɪtəl] 'little' by B7, and [ˈkʌpəl] 'couple' by B10. Other interesting observations include the fact that the youngest bilingual, B5, produced /l/ gliding in the word 'teletubbies' [ˈtejtʌbiː] in the Arabic session only and not in the English one (cf. Leopold, 1970) and that B7 produced a dark or emphatic /sˤ/ in the word 'muscle' [ˈmʌsˤl̩] where the /l/ nonetheless still sounded clear. Altogether, the English tokens that were produced in the Arabic sessions contained a mixture of features that belong to both languages and that were used within the 'same' words. Figure 3.36, 37, and 38 show illustrations for each of the bilinguals' varying production of the 'same' words in different language modes.

Although the code-switched tokens are not numerous, they definitely provide evidence for a different behaviour by the bilinguals in the production of the same variable depending on factors such as the base-language during which the production occurred and the interlocutor. More evidence will be provided in the following chapter from data on the /r/ variable.

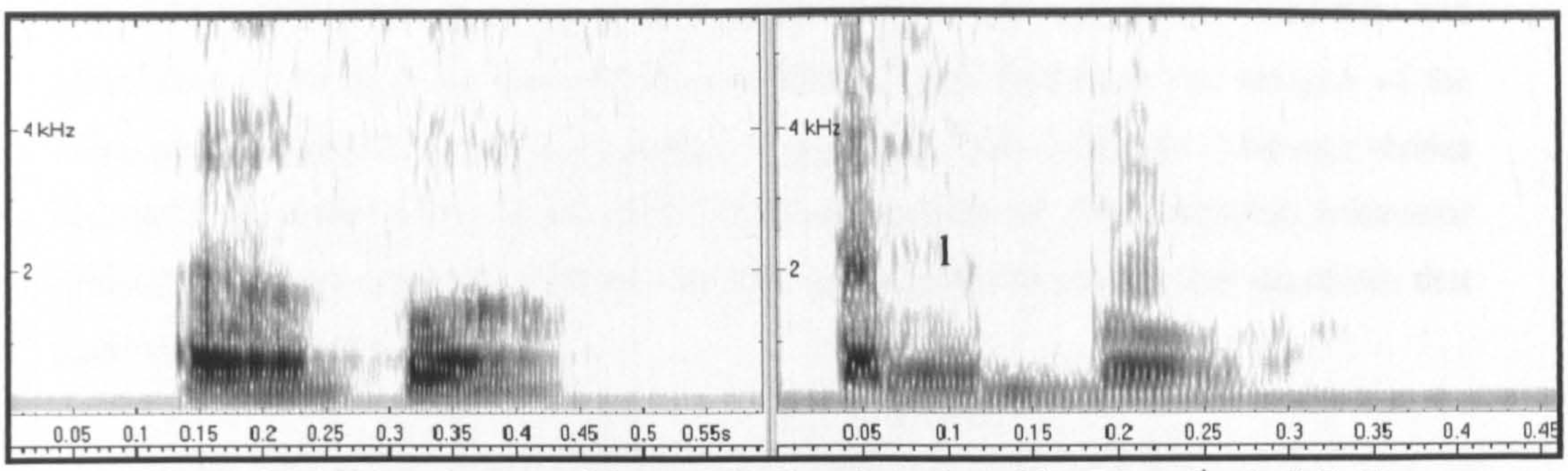


Figure 3.36: Spectrogram of the word ‘elbow’ produced by B5 as [ˈɛʊbə] (left) during the English sessions (F2 for [ʊ] = 11.19Z) and [ˈɛlbo] during the Arabic sessions (F2 for [l] = 13.74Z).

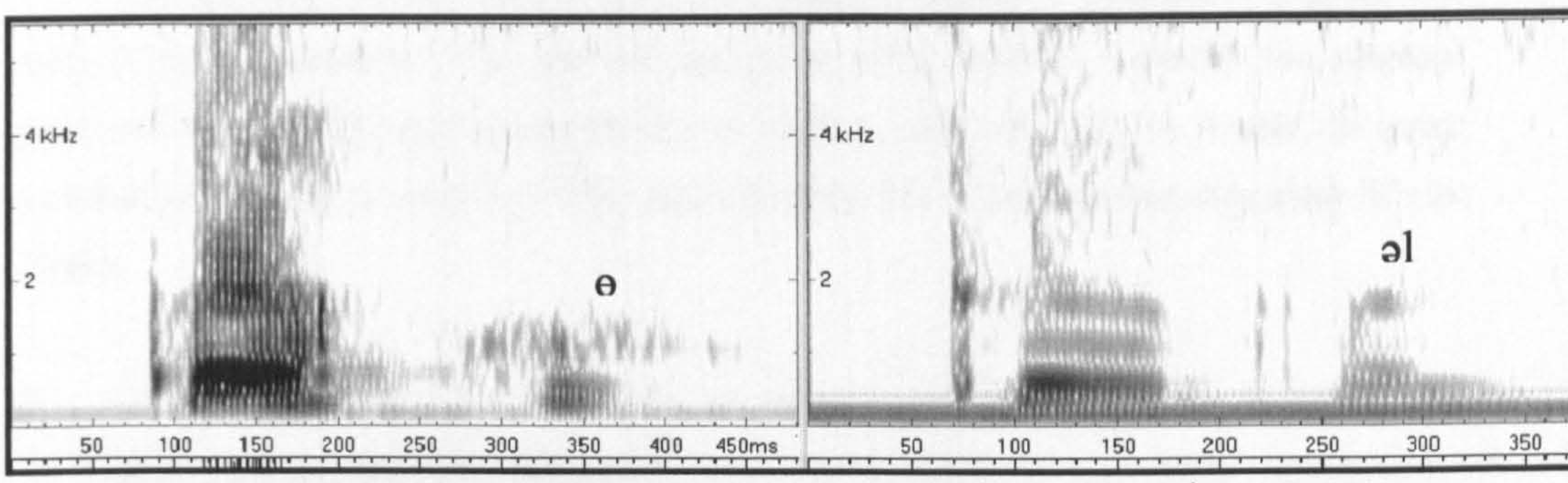


Figure 3.37: Spectrogram of the word ‘purple’ produced by B7 as [ˈpɜːpɵ] (left) during the English sessions (F2 for [ɵ] = 8.87Z) and [ˈpɜːpəl] during the Arabic sessions (F2 for [l] = 11.63Z).

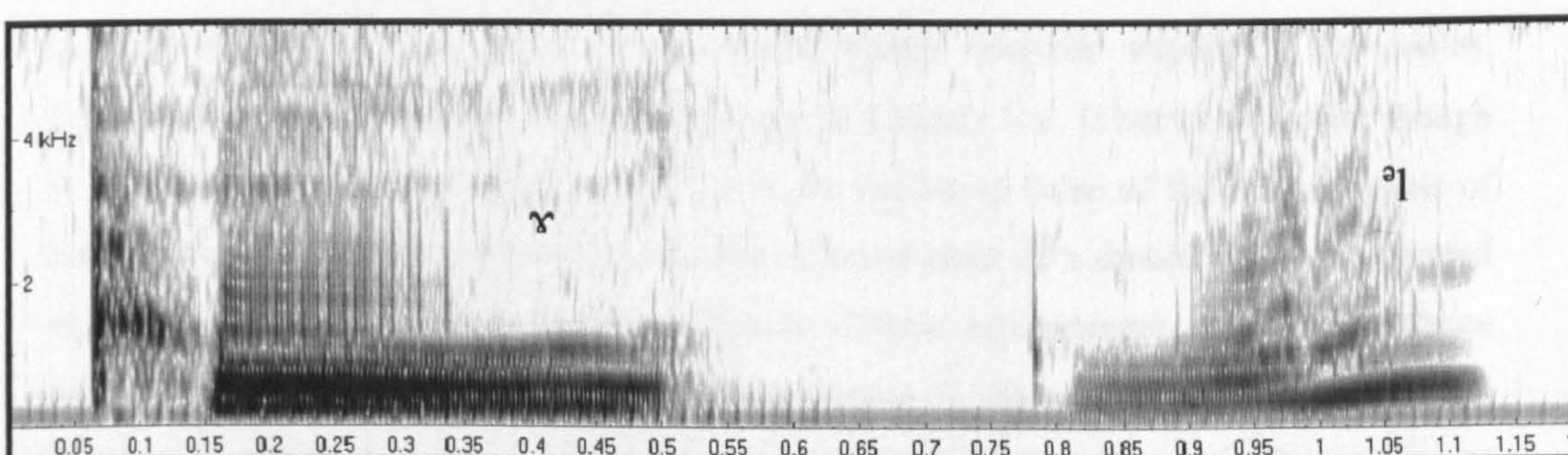


Figure 3.38: spectrogram of the word ‘pool’ produced by B10 as [pʰʉɤ] (left) during the English sessions (F2 for [ɤ] = 8.51Z) and [puːᵊl] during the Arabic sessions (F2 for [l] = 13.63Z).

3.10 Summary

The results obtained from this chapter offer important observations related to methodological issues in the study of bilingual phonological acquisition specifically and phonological acquisition in general. Although there are three bilingual children at the heart of the study, the inclusion of their parents along with monolingual children and adults from each language has offered a substantial contribution to the analysis of the bilinguals' production in the two languages. Moreover, if the bilinguals' language modes had not been taken into account, some misinterpretations of their linguistic behaviour might have been reached. An attempt will now be made to answer the four questions that were raised in Section 3.5.

1 Do English-Arabic bilinguals acquire separate /l/ production patterns for each of their languages?

The bilinguals in this study did indeed acquire separate /l/ production patterns for each of their languages. This showed mainly in coda position, whereby the subjects produced mainly dark and vocalised /l/'s in English, and clear /l/'s in Arabic. In onset position, the bilinguals produced both clear and dark /l/'s in English, but only clear /l/'s in Arabic.

2 Are their patterns of production in each language similar to those of the monolingual controls in the study?

The patterns produced by the bilinguals were on the whole similar to those of the monolinguals. In English, both groups of children produced few dark [ɫ]'s in onset position compared with some of the monolinguals' parents, especially the males. Explanations for this finding will be attempted in Chapter Six. What is important though is that results for the bilinguals turned out to be similar to those of the monolinguals of the same age, and therefore their production of initial clear [l]'s should not be interpreted as a failure to produce sociolinguistic aspects of their environment, or as an influence from their parents' L2 productions. The latter option is more obviously ruled out when looking at results for /l/ in coda position, since the bilinguals' parents show signs of language interference by producing a substantial amount of clear [l]'s, which are permissible in Arabic in this context.

The bilingual and monolingual children, on the other hand, produced dark final [ɫ]'s and a high number of vocalisations which seemed to increase with age, which rules out the possibility that it is simply a developmental feature. Since vocalisation was also found in the monolingual parents' productions, both groups of children may be showing signs of having acquired an accent feature that is available in their community. Moreover, data from recordings of the bilingual children in free-play sessions together and with their monolingual English friends produced similar results for the /l/ patterns found from the controlled sessions which I conducted. Since the free-play sessions consist of near-natural data, they provide strong support to the overall findings from this chapter and rule out the possibility that the bilinguals might have behaved differently during the sessions I conducted.

One minor difference between the two groups was noted in the small number of final clear [l]'s that the bilinguals produced; this may be due to influence from Arabic, since final /l/'s in Arabic are clear. Note, however, that the number of final clear [l]'s produced by the bilinguals in English decreases with age and that the youngest monolingual English child also produced sporadic clear [l]'s. Moreover, clear [l]'s in coda position only occurred during running speech, which underlines the importance of looking at several speech styles when analysing linguistic behaviour.

In Arabic, both groups of children produced clear [l]'s in onset and coda positions. As for emphatic contexts, /l/ realisations proved to be highly variable in their production due to developmental, contextual, social, and accentual factors (Section 3.2). As a result of this variability, it is not surprising that both the monolingual and the bilingual children produced a great number of clear [l]'s. What is interesting though is that two of the bilinguals (B7 and B10) actually produced more emphatic [lˤ]'s than the monolingual controls (A7 and A10). This was due to an accent feature that is only present in the speech of the two bilinguals' parents: BF7 and BM7 produced a number of emphatic glottal stops ([ʔˤ]) that are otherwise plain in the other bilingual and monolingual parents' speech; this increased the potential contexts for emphatic [lˤ], since emphasis in /l/ is mainly due to coarticulation. B7 and B10 have therefore acquired a feature of their parents' accent, which shows that they have the same ability as monolinguals to acquire emphasis given sufficient input.

One difference that was found between the bilingual and monolingual Arabic children is in the realisation of clear [l]'s in Arabic. An auditory and acoustic comparison of initial clear [l]'s in English and Arabic as produced by the monolingual children and adult showed subtle differences that need further investigation: Arabic [l]'s were found to be clearer than English ones (F2 frequencies for English clear [l]'s were lower than Arabic clear [ls] in comparable environments). The bilingual children, however, had no significant difference between the F2 frequencies for their English and Arabic clear [l]'s; in both languages, the F2 frequencies were closer to those of monolingual English patterns than to Arabic ones. This suggests that the bilinguals may have similar articulatory strategies for their English and Arabic clear [l] productions, but that these strategies are different from those of the monolingual Arabic children.

3 Are the patterns for the monolingual subjects in this study similar to the ones normally described in the literature and therefore expected for each language?

In English, although it was expected that Yorkshire /l/'s would be dark-ish in all positions (Wells, 1982: 371), data from the Leeds IViE corpus and the monolinguals' parents suggested otherwise. It emerged that some but not all of the adults from the IViE corpus and the monolinguals' parents recorded for this study produced initial dark [ɫ]'s. Possible reasons for this outcome are gender- and accent-related and will be discussed in Chapter Six. As for coda /l/'s, all speakers produced the expected dark variety, and most speakers also produced /l/ vocalisations, which varied in frequency from one speaker to another.

In Arabic, results for the monolingual parents and their children confirm the fact that Arabic /l/ is clear in all contexts except emphatic ones. An examination of the patterns found for /l/ in emphatic contexts also confirms the huge amount of variability in its production that is reported in the literature (e.g. Kahn, 1975; Lehn, 1963; Mitchell, 1993). Such variability is due to contextual, dialectal, social, and developmental factors and points to the difficulty in assessing whether any child has acquired its production. The adult results also revealed the existence of an accent feature in two of the bilinguals' parents that is not well-documented in the literature and that involves an emphatic production of the glottal stop ([ʔˤ]) in contexts where the glottal stop historically originates from a uvular plosive ([ħalˤaʔˤ] for [ħalaʔ], historically /ħalaq/). The availability of this context resulted in more emphatic productions of /l/ for two of the bilingual children in the study compared with the monolingual controls.

4 Are there signs of influence from one language onto the other in the bilinguals' production and what are the factors that affect such influence?

Two types of influence are noted here: the first one concerns the small number of clear [l]'s that were produced by the bilinguals in coda position in English, and the lack of distinction between English and Arabic clear [l]'s compared with the apparent subtle distinction in F2 that was found between the monolingual English and Arabic controls (though this issue definitely needs a more controlled experiment to verify the results). This first type of influence was minimal and did not show a great deal of interaction between the two languages.

The second type of influence concerns the bilinguals' English productions during the Arabic sessions. The /l/ productions from these sessions were examined separately from the ones that were produced during the English-only sessions due to the effect of the language mode on the variants chosen by the bilinguals. This study provides support for the concept of language mode from a phonological point of view, as the English /l/ variants that the children produced during the English-only sessions differed qualitatively from those produced during the Arabic sessions. Explanations for this behaviour will be attempted in Chapter Six.

CHAPTER FOUR

/r/ production

4.0 Introduction

This chapter presents results from auditory and acoustic analysis of /r/ production in English and Arabic. In Section 4.1, /r/ production in English is described, taking into account the variety of English /r/ produced in the bilingual subjects' environment and developmental patterns of /r/ acquisition normally found in children. Section 4.2 offers a similar description for /r/ production in Arabic, and Sections 4.3 and 4.4 present what is known about bilingual acquisition of /r/ and sociolinguistic factors that may affect such acquisition. The aims for this chapter are then listed in Section 4.5, followed by a description of the material used for /r/ examination and the type of analysis conducted in Section 4.6. The detailed results for the subjects are then presented in Sections 4.7 and 4.8, and a summary of the main patterns follows in Section 4.9, along with a discussion how the findings of relate to the aims of this chapter.

4.1 English /r/

4.1.1 Articulatory description of English /r/

In most English accents, /r/ is produced as a voiced alveolar or post-alveolar approximant [ɹ], although the tap [ɾ] remains the localised variant found in many parts of northern England, Scotland and Wales (Hughes & Trudgill, 1996: 90; Wells, 1982: 368). In the production of the approximant, the tongue tip is held in a position near the rear part of the alveolar ridge, the back rims of the tongue are touching the upper molars, and the central part of the tongue is lowered, creating some sulcalisation (grooving) behind the tip/blade stricture. There is also a general retraction of the tongue creating a pharyngeal constriction and an effect of hollowing and slight retroflexion of the tip, normally accompanied by lip rounding (Cruttenden, 2001: 206; Jones, 1972: 195; Ladefoged & Maddieson, 1996: 233; O'Connor, 1973: 150; Roach, 1991: 60). The degree of labialisation varies across speakers, some of whom labialise /r/ whatever the following vowel (Cruttenden, 2001: 207). Another /r/ variant is a labiodental one that is normally transcribed as [ʋ], and that is increasingly becoming a feature of many urban English accents (Foulkes & Docherty, 1999, 2000; Hughes & Trudgill, 1996; Kerswill, 1996; Williams & Kerswill, 1999; Llamas, 1998; Stuart-Smith, 1999; Trudgill, 1999). Though in the past [ʋ] was originally thought to be a feature of immature or defect speech, its emergence as a non-standard accent feature in the South of England and the change in its perception may have contributed to its spread to other English varieties as it now seems to be growing as a variant for young working class people across urban areas of Britain (cf. Foulkes & Docherty, 1999 for a comprehensive discussion of the origin of [ʋ] and the rise in its status). [ʋ] is principally characterised by a labial articulation, although results from acoustic and visual analysis suggest that it may lack both lower lip retraction and lingual articulation (Foulkes & Docherty, 1999). The phonetic characteristics of [ʋ] are actually quite variable across speakers and within the speech of individuals.

There is no detailed description of the /r/ variety used in Yorkshire, and the reports available in the literature are either too general or outdated. Some of the available accounts of parts of Yorkshire include Hughes & Trudgill (1996: 90), who describe the /r/ in Bradford as a flap, and Stoddart et al (1999:76), who describe the Sheffield /r/ as mainly an approximant although a tap can be heard among males. Wells (1982: 368) notes that the alveolar tap [ɾ] seems quite widespread in the north of England as a rival to the usual post-alveolar approximant and associates its use with Leeds, but also admits that the geographical spread of the tap is not well known.

4.1.2 Phonotactic and phonological distribution of /r/

In English, /r/ has a restricted distribution. In syllable onsets, /r/ must occur adjacent to the nucleus. Where the onset consists of a consonant cluster, /r/ can cluster with nine different obstruents: /p/, /b/, /t/ /d/, /k/, /g/, /f/, /θ/, and /ʃ/ (Cruttenden, 2001: 201). For rhotic accents, /r/ in syllable codas must also occur next to the nucleus. For historical reasons, however, post-vocalic /r/ is absent before a consonant or in absolute final position in several accents of English (*farm* [fɑ:m]; *far* [fɑ:]) (Cruttenden, 1994: 268; Foulkes, 1997b: 260; Hughes & Trudgill, 1997: 60; Wells, 1982: 218). One of the interpretations offered on the historical change is that the English /r/ went through several stages (from a trill/tap, to a fricative continuant, and then to frictionless approximant) before it merged with a preceding vowel in final or pre-consonantal positions. Post-vocalic /r/ later became regionally restricted, as some dialects but not others preserved it as an accent feature (Cruttenden, 1994: 189).

Although most English English dialects are non-rhotic, there is no clear description of the variety used in Yorkshire. Both Wells (1982: 368) and Hughes & Trudgill (1996: 33) note that most urban Yorkshire accents are non-rhotic even though some of the traditional rural areas in East Yorkshire are still characterised by a partial retention of post-vocalic /r/. There was so sign of rhoticity in the data analysed from the Leeds IViE corpus (see Section 4.7).

Even in non-rhotic accents, post-vocalic /r/'s still appear in certain environments and according to certain rules. There are debates as to whether or not the /r/'s in this position are encoded in the lexical representation of words; for this reason, post-vocalic /r/ realisation is explained in terms of either insertion or deletion rules, or both (see discussion in Foulkes, 1997b: 270). The emerging patterns are usually described as:

- Linking R: where the /r/ was historically present in pre-pausal position, and the next word following /r/ begins with a vowel, the /r/ may be pronounced and the feature is known as 'linking' /r/ (*brothe[r] in law*).
- Intrusive R: where the [r] realisation is introduced into words which did not historically contain an /r/, it is labelled as 'intrusive'. This phenomenon is found in certain intervocalic environments (Foulkes, 1997b: 260; Wells, 1982: 224):

$$\emptyset \rightarrow r \,/\, [\text{ɔ:, ɑ:, ə}]\,\_\,\#\,V.$$

e.g.: *Is Ma /r/ at home?* *law /r/ and order* *withdraw /r/ al.*

Unlike linking /r/, intrusive /r/ is often regarded as incorrect (Wells, 1982: 224; Trudgill, 1974: 162). However, the view is not so straightforward; a recent investigation in Newcastle has found evidence of a relationship between the use of both processes and social as well as stylistic factors (Foulkes, 1997b).

4.1.3 Acoustic description of English /r/

The English approximant is normally characterised by a weak vowel-like acoustic structure made up of a series of weak formants due to a narrower constriction than that normally made for vowels. This is manifested by an average of 10dB lower amplitude in the liquid compared with a following stressed vowel (Stevens, 1998: 534). F1 is normally between 120 and 600Hz, with the lower frequency giving greater impression of lip rounding, while F2 is between 700 and 1200Hz, with the lower frequencies tending to occur in initial position and the higher ones in intervocalic position. The retroflex and rounded variant of [ɹ] is distinguished by a particularly low F3 that is close to F2 (Borden & Harris, 1984: 113; Cruttenden, 2001: 207; Ladefoged & Maddieson, 1996: 234; Stevens, 1998: 535[6]), while energy above F3 is normally very weak due to the existence of two anterior constrictions in the vocal tract, one made by the tongue tip or blade, and the other by the narrowed lips. F3 for [ɹ] ranges widely between 1400 and 2400HZ

6 Stevens argues that the resonance that appears close to F2 is not actually F3, but a new resonance that he calls F_R created by the front cavity anterior to the point of constriction (caused by the side chamber under the tongue tip retroflexion). Still, the effect of F3 or F_R is that of a dipped energy into the approximant and a sharp rise out of it.

depending on individual, gender, and contextual variation (Foulkes & Docherty, 2000: 50; Nolan, 1983: 93). For instance, in his study of English liquids as produced by 13 adult male speakers of Southern British English, Nolan (1978: 30) reports mean frequency values of 320Hz for F1, 1090Hz for F2, and 1670Hz for F3 in initial position. With respect to contextual variation, F3 for [ɹ] tends to be lower in initial than in intervocalic position (Foulkes & Docherty, 2000: 50).

Due to its weak formants, [ɹ] tends to be identified on spectrograms by its steeply rising transitions to a following vowel. As for the labiodental, there are few examples of acoustic analysis of it in general, and only one study with acoustic description of the variant occurring in British English (Foulkes & Docherty, 2000). Contrary to [ɹ], [ʋ] displays little phonologically-conditioned variation in F3 and is generally characterised by a dip in all formants, though its F3 is still higher than that of [ɹ], with an average of around 200 Hz difference. Foulkes & Docherty note that the possible lack of F3 is a sign of little or absence of tongue retroflexion or bunching that is typically associated with [ɹ].

4.1.4 Acquisition of /r/ by monolingual speakers

As with /l/, few studies have focused specifically on the development of /r/, and the information gathered here for its acquisition is taken from more general studies of phonological development (Bernhardt & Stemberger, 1998: 305/331; Cruttenden, 2001: 209; Edwards, 1973: 9; Ingram, 1979: 135-140; Matthews, 2001: 216-218; Menyuk, 1971: 80; Moskowitz, 1970; Sander, 1972: 62; Smith; 1973: 2/18/75; Vihman, 1996: 219/ 239). In English, the production of liquids emerges relatively late, preceded by early production of nasals, plosives, and some of the fricatives. The production of [ɹ] is known to involve physically complex articulations and usually emerges later than /l/ in children's speech, commonly around the age of 4;5. [ɹ] production is highly variable and is not normally mastered before the age of 6, with mature production in prevocalic contexts generally preceding that of post-vocalic ones. [ɹ] is frequently replaced by [w] and [ʋ] in initial position, e.g. 'rabbit' [wæbɪʔ]; 'red' [ʋɛd], and less commonly by [l] and [j] e.g. 'rain' [leːn]; 'room' [juːm]. /r/ is often deleted in initial consonant clusters and in medial and final position (for rhotic accents), e.g. 'grandma' [ˈgɛːmɑ]; 'dress' [dɛs]; 'very' [vɛiː]; 'car' [kɑː]. In initial clusters with alveolar stops, stops are often affricated or fricated, e.g. [dʒɛs]. Another process common in early productions is stopping e.g. 'rat' [dæt]; 'record' [gɛˈkɔːd].

4.2 Arabic /r/

4.2.1 Articulatory description of Arabic /r/

Arabic /r/ is normally a tap or a trill, depending on free and allophonic variation (Anani, 1985: 132; Nasr, 1966: 5; Shaheen, 1979: 142). Allophonic variation is mainly concerned with the distinction between single and geminate /r/'s in intervocalic position, whereby single /r/'s are produced as taps (['baɾa] 'he sharpened') and geminates as trills (['barra] 'outside'). There is however, free and individual variation in the production of taps and trills, so that single /r/'s are sometimes trilled while geminates are produced with a single long tap. Also, like any other language where a possible realisation is a trill, not all speakers use a trill and even those who do use trills have taps and other realisations as well (Lindau, 1985: 161).

In the production of the alveolar trill, the tongue blade or tip is brought into complete closure with the alveolar ridge using a light contact pressure, which allows the oral pressure to build rapidly and force its way through the closure. Then a combination of elastic muscle forces and the sucking action of the Bernoulli effect bring the tongue back into renewed contact, thus creating repeated cycles (Catford, 1988: 69; Laver, 1994: 219). There is, however, the possibility of a trill being produced by a single pulse or closure followed by a prolonged opening phase rather than several pulses (Lavoie, 2001: 83; Lindau, 1985: 161). Such a description can help explain how geminate /r/'s in Arabic sometimes sound like long taps rather than trills. Lavoie (2001: 143-144) further notes that the number of pulses per trill may vary according to context; her findings on Spanish trills show greater number of pulses for trills that are produced in stressed rather than unstressed positions.

In the production of the tap, the tip of the tongue typically makes contact with the alveolar ridge, while the back cavity is characterised by a wide unobstructed pharynx and a gradual narrowing towards the region of the articulatory constriction. There is, however, slight disagreement in the literature concerning the nature of the tap and the difference between it and the stop. This might be due to the fact that taps are produced differently depending on context (Lavoie, 2001: 84), language, and even speakers of the same language (Lindau, 1985: 161). In most accounts, the tap is considered to involve a very fast movement of the tongue in the onset phase, an extremely brief closure period, and a very fast offset, making it altogether much shorter than a stop in the same place of articulation e.g. [d] *versus* [ɾ] (e.g. Catford, 1988: 71; Laver, 1994: 224). Recent electropalatographic evidence from English, however, shows that the tap is similar to a stop in duration, especially in onset and closure, but that the degree of linguo-palatal contact in the tap is smaller and the closure is often incomplete (Connell, 1995: 43).

Connell further notes that the offset phase of the tap is actually longer than that of the stop, suggesting a weaker gesture in the production of the tap. Lavoie (2001: 84) goes as far as placing the Spanish tap under the approximant category when it occurs in medial position.

Descriptions of the Arabic tap echo the diversity of realizations found in other languages. Shaheen (1979: 142-145), for instance, notes that although Arabic /r/ is always labelled as a tap (or a trill), it can be phonetically realised as a tap, a frictionless continuant or a fricative. The symbols used by Shaheen for the three variants of each category are [ɾ], [ɹ] and [ɹ̝]. Shaheen further notes that the position of /r/ in the word has a considerable effect on its spectrum and its duration, which ranges from 25 to 180 ms (shortest in intervocalic position and longest in final position), which also fits in with Connell's suggestion of a weaker gesture in the tap than in the stop. This study will offer further support for Shaheen's claim that the Arabic tap can be realised as a continuant, but will suggest [ɾ] rather than [ɹ] as the symbol for it, mainly due to marked phonetic (auditory and acoustic) differences between the realisations in each language, although the resulting variant in each case is that of an approximant quality.

4.2.2 Phonotactic and phonological distribution of Arabic /r/

Similarly to English, Arabic /r/ occurs adjacent to the nucleus in both syllable onsets and codas. However, as opposed to the restricted context in which English /r/ can occur, Arabic /r/ can cluster with more obstruents than English, including /b/, /d/, /dˤ/, /t/, /tˤ/, /k/, /ʔ/, /f/, /s/, /z/, /ʃ/, and /x/, as well as clustering with other sonorants like /m/ and /n/.

As opposed to the absence of post-vocalic /r/ in non-rhotic English accents, Arabic /r/ is produced in all pre- and post-vocalic contexts. Moreover, Arabic /r/ can occur as the nucleus of initial e.g. [ɾ̩ˈbɪħna] 'we won' and final syllables e.g. [ʔabɾ̩] 'grave', and also subject to gemination e.g. [ˈbarːa] 'outside'.

4.2.3 Acoustic analysis of Arabic /r/

Since the production of taps and trills is characterised by one or several rapid interruptions of the air stream, their spectra typically have similar acoustic features to plosives along with a vowel-like formant structure and/or friction-type noise that are visible between the short gaps (Ladefoged & Maddieson, 1996: 218; Shaheen, 1979: 142).

In initial position, descriptions of the Arabic tap mention the presence of distinct formant structures interrupted by a short vertical gap with a duration of around 15-20ms, while the trill is characterised by multiple vertical gaps and can be acoustically regarded as a series of taps (Al-Ani, 1970: 33; Lindau, 1985: 166; Shaheen, 1979: 145-160). Acoustic energy is concentrated in well-defined formants only at the lower end of the spectrum. In 80% of /r/ occurrences F3 is absent and acoustic energy above F2 is unevenly distributed, though vaguely anticipating the formants of the following vowel. In the absence of a gap, /r/ is said to appear as a frictionless continuant [ɹ]. The average steady-state position of F1 is 305 Hz, that of F2 is 1310 Hz, while F3 when present is around 2400 Hz. It is interesting to note that while for English /ɹ/ F2 and F3 are very close, F2 and F3 (when present) in Arabic are often widely separated (Shaheen, 1979: 145-160).

In intervocalic position, the spectrum of /r/ is described as being similar to that of a stop. It appears on the spectrogram as a gap with no energy above the voice bar, apart from occasional appearance of a shadow of the formants of the adjacent vowel. Intervocalic /r/ has an average duration of 25ms (Shaheen, 1979: 145-160).

As for final position, when devoiced, /r/ can often show a spectrum of a fricative nature [ɹ̊]. In half of the occurrences of [ɹ̊], acoustic energy is diffusely spread in the frequency range 2700-5000Hz, while in the other half of the occurrences, although acoustic energy is still diffusely spread among the frequencies, F1 and F2 could be detected despite their low intensity. The average steady-state frequency of F1 is 250Hz, while that of F2 is 1420Hz. It is interesting to note that for Arabic, as opposed to what is normally reported for English, F2 can be higher in final position than in initial position, while the opposite applies for F1 (Shaheen, 1979: 145-160).

Keeping in mind the differences described above between English and Arabic /r/, the discussion now moves to a frequently researched question in bilingual acquisition, that of whether bilinguals acquire similar or separate patterns for their production in each language, before we move back to the discussion of the production of /r/ by the subjects in this study.

4.2.4 Acquisition of /r/ by monolingual speakers

Similarly to English, /r/ production in Arabic is usually more difficult to acquire than /l/ and may be replaced by /l/ in the initial stages (Amayreh & Dyson, 1998: 646; Omar, 1973: 48-56). Dyson & Amayreh, (2000: 84) actually group /r/ under the most difficult sounds to acquire along with the emphatics due to the articulatory complexity that is involved in its production. Though Arabic /r/ emerges around the age of 3, it only reaches an acceptable performance towards around the age of 5;6, approximately the same age as that of the acquisition of the English /r/.

Monolingual developmental features for Arabic /r/ normally include deletion e.g. [naː] 'fire' for adult [naːɾ], assimilation, e.g. [ʔikkab] 'I ride' for adult ['ʔirkab] and substitution, which being more frequent and mainly involves lateralisation, e.g. ['lasam] 'he drew' for adult ['ɾasam]. Lateralisation shows a clear developmental trend, declining rapidly from early production till the age of 4;4, and normally disappearing after 5;5 (Dyson & Amayreh, 2000: 89-91; Omar, 1973: 56). Another occasional type of /r/ substitution is gliding of /r/ to [j], but there are normally no occurrences of [w] for either /r/ or /l/. This pattern is quite different from English where /r/ gliding to [j] but mainly [w] is frequent whereas lateralisation is uncommon, although it does occur occasionally (Smith, 1973: 75). This may be due to the fact that [ɾ] and [l] share tongue tip contact, while [ɹ] and [l] don't, while [ɹ] and [w] involve labiality. A final rare type of /r/ substitution reported in Dyson & Amayreh (2000: 94) is stopping, though there is no mention of the stop variants produced.

4.3 Bilingual acquisition of /r/

Acquisition of /r/ by bilinguals has only been looked at in the early stages of acquisition and as part of case studies of the overall bilingual phonological development of a given child e.g. Burling (1971), Leopold (1970), and Ingram (1982). Each of these studies will be discussed briefly in this section. The only study I am aware of that is dedicated to bilingual acquisition of /r/'s in particular is that by Ball, Muller & Munro (2001b) and will be discussed in greater detail towards the end of this section.

In Leopold's (1970) longitudinal study of his English-German bilingual daughter's production (also discussed in Chapter Three), the author notes that Hildegard did not produce any /r/'s during the first two years of life. Leopold attributes that to the difficulty in articulation of the English /r/ for children, mainly 'the complicated adjustment of the tongue muscles for raised position of the tongue tip' (Leopold, 1970: 64), and the fact that English and German /r/'s are so different, the German /r/ being a velar fricative [ɣ]. As opposed to /l/, which showed early signs of separation in Hildegard's production in each language, Leopold (1970: 64) notes that Hildegard still treated the German and English /r/'s in the same way in terms of omissions and substitutions. For instance, initial /r/ was constantly replaced by [w], which Leopold explains as serving the labial nature of the English /r/ and the raised tongue back position of the German velar. Final /r/'s were omitted or substituted by vowels of varying quality, but lacked labialisation. One such substitution was [ə], which is usual in North German colloquial pronunciation Leopold (1970).

Burling's (1971) study (also discussed in Chapter Three), describes a case of early differentiation between the patterns of liquid production by his English-Garo speaking child between the ages of 1;4 and 2;8. Although the author notes that his son's awareness of the two languages being different only emerged at the age of 2;2, the description of earlier productions of /l/'s and /r/'s in the two languages provides evidence for differentiation. For instance, between the ages of 1;5 and 2;8, Stephen used [l] for both Garo [l] and [ɾ] e.g. [lama] for /ɾama/ 'road' (in Garo the two sounds are allophones of the same phoneme, with [ɾ] occurring in syllable-initial position and a lateral similar to English [l] occurring elsewhere), while he replaced English /r/ with [w] or omitted it altogether. None of the Garo liquids were replaced with the labial-velar approximant.

A similar observation is noted by Ingram (1982) in his study of his Italian-English bilingual daughter, which provides another piece of evidence for two different phonological patterns for /r/ as produced by the child. Similarly to Stephen (Burling, 1971), Ingram's daughter Jennika substituted [l] for /r/ in Italian (e.g. [lakonta] for [ɾakonta] 'story') and [w] for /r/ in English (e.g. 'ready' [wedi]).

As mentioned before, Ball et al's study (2001b) is the only extensive analysis of the bilingual acquisition of /r/ and will therefore be discussed in more detail. The type of /r/ variants that Welsh-English bilinguals need to acquire is similar to that of the English-Arabic bilinguals in this study. The authors examined the developmental patterns in the acquisition of rhotic consonants by 85 Welsh-English bilingual children between the ages of 2;6 and 5;0, divided into five age ranges and into Welsh-dominant or English-dominant subjects. Since Welsh is spoken by about half a million speakers in Wales, the authors managed to examine subjects who were mastering both languages simultaneously in a predominantly bilingual environment. Welsh has both a voiced and a voiceless alveolar trill [r] and [r̥ʰ] which occur in all word-positions, whereas the accent of the English spoken by most Welsh-English bilinguals is (mainly) non-rhotic and uses a post-alveolar approximant [ɹ] (Ball et al, 2001b: 72).

The study confirms the difficulty in the acquisition of rhotics in that the Welsh trill was acquired with only 50% accuracy by the oldest Welsh-dominant bilinguals, whereas the English approximant was acquired with only 30% accuracy by the oldest English-dominant bilinguals (Ball et al, 2001b: 73). More intriguing was the varied number and quality of substitutions that the subjects exhibited in their /r/ productions in both languages, particularly for the Welsh trill but for the English approximant as well. Realisations other than the trill included other approximants or liquids, nasals, fricatives, stops, consonant clusters, and deletions (Ball et al, 2001b: 74).

Another interesting finding in Ball et al's study is that the rate of acquisition differed between the two dominant groups and decreased over time. For the English-dominant subjects speaking Welsh, the prevailing trend was to use the approximant [ɹ] in both initial and medial position, with few instances of substitutions after age 4;0, while deletions in word-final position were very frequent till age 4;6. The authors interpreted the deletions in terms of English influence and the fact that very few Welsh-English speakers use postvocalic /r/. Substitutions in word-final position included fricatives, liquids, and glides in most groups, and trills only by the oldest group of speakers (4;6-5;0). The Welsh-dominant bilinguals speaking Welsh differed from the English-dominant subjects in their higher use of the target form [r] in initial and medial position, but also in the greater variability of substitutions, which the authors interpreted as a faster developing system. In word-final position, however, the patterns that were found were similar to those of the English-dominant bilinguals, mainly including trills, but also deletions, [ɹ], fricatives, liquids, and glides. The authors interpreted the use of [ɹ] for the trill as both an articulatory strategy to avoid the complex articulations involved in the trill, but also as interference from English, since the same subjects who used [ɹ] in Welsh were still acquiring the English variant and using other substitutions for it. The use of variants like fricatives and liquids for the trill by both English- and Welsh-dominant groups was taken as articulatory and acoustic strategies to achieve the noise and continuance components of the trill.

With respect to the acquisition of the English [ɹ], the degree of variability in both dominance groups was much less than in their Welsh production, and substitutions were restricted to liquids and approximants, with few deletions and fricative or nasal variants [v], [ʁ] and [n], and higher correct percentage rates in the post 5;0 group (Ball et al, 2001b: 79). More importantly, there was no final /r/ production in any of the dominance groups (apart from a few linking-r instances), which conforms to the non-rhotic variety present in the subjects' Welsh-English accent. There was also a predominance of the [ʋ] variant, which the authors interpreted as an acoustically-driven process due to the similarities in both duration and formant structure of [ʋ] and [ɹ]. Overall, Ball et al's study showed that differences in rate of acquisition and amount of variability are clearly linked to the dominant language of the subjects, and the use of substitutions derives from acoustic as well as articulatory similarity with the target sound.

So far, all the studies mentioned concentrate on the early stage of bilingual acquisition, and the results are mixed with regards to whether or not the children show early signs of differentiation in their production of the relevant language-specific /r/

variants. What is missing, though, is data from the later stages of the bilinguals' development in order to examine whether or not they fully separate their production of /r/'s in each language. Though most research on bilingualism has concentrated on early evidence for separation, not much is mentioned about later stages of bilingual development, and the overall view is that all bilinguals eventually develop separate systems. More detail is needed on the state of bilingual children's production in the later stages of development. Also, another important issue that has not been discussed by any of the studies above is the influence of the language mode on the type of variants that the bilinguals might choose to use. Finally, there is barely any research on how English-Arabic bilingual children acquire the patterns for /r/ in either language, although there are important differences with respect to its place and manner of articulation and in the phonotactic constraints and phonological patterning that govern its production and occurrence/ distribution in each of the languages.

4.4 Sociolinguistic issues in the acquisition of /r/

As mentioned in Chapter Three, few studies have considered the phonological repertoire of bilingual children with the particular local accent(s) spoken in their environment in mind in order to examine the motivating factors that instigate the production of one realisation over a number of competing alternatives.

In Verma et al's (1992) study (reviewed in Chapter Three), the Panjabi/ Urdu speaking subjects from West Yorkshire seemed to alternate between the tap [ɾ] and the trill [r], which the authors interpreted as a combination of influence from their L1 and their local variety (Verma et al, 1992: 189). It is interesting to note, however, that the children did not produce the retroflex variant [ɽ], which would have been a clear influence from their L1. The subjects did have a rhotic accent, and therefore tended to produce postvocalic /r/'s as in 'star' [staɹ], and 'water' [wɔʔər]. Though Verma et al interpret this as transfer from the subjects' mother tongue in which all orthographic /r/'s are produced, the subjects have learned their English primarily from school, and English orthography may have made it more difficult for them to acquire non-rhoticity.

Similarly, Agnihotri (1979) examined processes of assimilation to Leeds English that Sikh children of immigrant families exhibit in relation to their length of stay in Leeds and found an overall negative correlation between 'accent-revealing' features (features that would identify them as non-native speakers) and length of stay in Britain. There were, however, other important determining factors such as speech style, gender, social background and area of residence of the families involved. For instance, the occurrence of post-vocalic /r/ tended to decrease in the subjects' production not only with the length of stay, but also in casual style as opposed to reading style, with females more than males, for Sikh children of Indian origin rather than Kenyan origin, and for Sikh children who interacted more frequently with native Leeds English speakers than those who lived in immigrant areas (cf. Agnihotri, 1979: 243-253 for a discussion).

More interestingly, Agnihotri found that the English of each of the individuals showed simultaneous use of features from the different varieties that they were exposed to from native and non-native speakers of the language. This mixed code exhibited itself in the way the children produced the same sounds sometimes 'the Indian way' and other times 'the English way' in the same utterance. For instance, the author gives the example 'mother', which was produced by the children as [mʊðə], [mʊðər], or even [mʊðəɹ], sometimes within the same utterance. The use of features from all varieties by the young bilingual has also been discussed by Heselwood & McChrystal (2000), although Agnihotri attributed their use to code-mixed utterances whereas Heselwood & McChrystal found such features even in their attempt to elicit data from their subjects in a monolingual mode. Such results are very important because they highlight the complexity of factors that interact in shaping the acquisition of young learners and the necessity of looking at individual as well as group results when interpreting data from bilinguals.

4.5 Aims of the study

In light of the preceding discussion which has drawn attention to the importance of taking social dimensions into consideration when defining a 'phonological system' and of the role of the language mode in analysing bilingual data, this study examines the extent to which bilingual children can establish phonetically/phonologically distinct production patterns for /r/ in each language. The experiment is designed to investigate the following questions:

1. Do English-Arabic bilinguals acquire separate /r/ production patterns for each of their languages?
2. Are their patterns of production in each language similar to those of the monolingual controls in the study?
3. Are the patterns for the monolingual subjects in this study similar to the ones normally described in the literature and therefore expected for each language?
4. Are there signs of influence from one language on the other in the bilinguals' production? If so what are the factors that affect such influence and how are they related to the bilingual's language modes?

4.6 Procedure

4.6.1 Material collected for the /r/ experiment

As with /l/, data for this chapter are taken both from recordings from the Leeds IViE (Intonational Variation in English) corpus (Grabe & Nolan, 2001) and the recordings collected for this study. /r/ tokens from Grabe & Nolan's data were collected from three different speech styles (sentences, reading passage, and story telling). Around 100 /r/ tokens per speaker in a variety of vocalic contexts were then available for analysis.

Material from my own recordings was extracted from words produced in isolation during the picture-naming activities for the children and the reading lists for the adults, and in running speech from the story telling activities with both children and adults and interviews with the adults (Table 4.1). All the words that had 'r's in the spelling as well as in the pronunciation were examined in both languages in order to compare the occurrence of post-vocalic /r/'s by different subjects and in different languages.

Table 4.1: Sample tokens used for the examination of /r/ in English and Arabic

English	**pre-vocalic**		**postvocalic**	
Examples	giraffe orange red carrot		butterfly worm deer horse	
Arabic	**pre-vocalic**		**postvocalic**	
Examples	**IPA**	**Gloss**	**IPA**	**Gloss**
	zaɾa:fe	giraffe	ʔaħmaɾ	red
	bɪr:a:d	fridge	xja:ɾ	cucumber
	faɾa:ʃe	butterfly	bɪrnajtˤɑ	hat
	ʒazɾa	carrot	kɪrse	chair

4.6.2 Analysis

While auditory analysis was conducted on the IViE data and running speech collected for this study, both auditory and acoustic investigations were conducted on the words produced in isolation by the children and adults from the current study. With respect to the auditory analysis, the /r/ tokens that were produced were initially coded for one of ten categories (Table 4.2), including four choices for the obstruent-like type, four choices for the approximant type, one for deletions and a final one for other realizations. The decision behind this categorisation was made during the auditory analysis in order to avoid as much as possible forcing variants into rigid categories, without losing sight of the aim of the investigation, which in principle is to find out whether the bilinguals will produce language- and accent-specific /r/ variants.

Table 4.2: number of categories devised for labelling the /r/ tokens in English and Arabic

Stop-like types				Approximant types					
1	2	3	4	5	6	7	8	9	10
r	r or ɾ	ɾ	ɾ̞	ɹ	weak ɹ	weak ʋ	ʋ	Ø	other

With respect to the stop-like types, two categories were added to the conventional tap and trill types: (i) category 2 was chosen when it was not auditorily clear whether the token was a trill or a tap, partly because trills can be sometimes be realized as single taps (e.g Lavoie, 2001; Lindau, 1985), but also because some bursts are fainter and lower in amplitude than others and are therefore difficult to distinguish by ear alone; later acoustic investigation helped confirm the choice as either [r] or [ɾ] and category 2 was deleted (ii) category 4 was included when it was noticed that some of the productions of the Arabic tap involved no contact between the tongue tip/blade and the alveolar ridge and were heard more like approximants than taps. Although this type of weak stop realization has been noticed before and labeled [ɹ] by Shaheen (1976), a different label was necessary due to the fact that the realization still differed from the English approximant [ɹ] in that there was no audible reflex of retroflexion or lip rounding. For this reason, [ɾ̞] was chosen instead and later acoustic analysis offered further evidence for its nature and its distinction from [ɹ]. The distinction between [ɾ̞] and [ɹ] proved important in the analysis of the bilinguals' production in the two languages (see Sections 4.7 and 4.8).

As for the approximant types, a 4-point scale was devised once again to allow for the variation that was found in some speakers' pronunciation rather than forcing tokens into the alveolar or the labial variety (cf. Foulkes & Docherty, 2000). Finally, 'Ø' and 'other' realizations have categories of their own because of the importance of the occurrence of each one in terms of contextual, developmental, and sociolinguistic differences. For instance, in the English data, all the tokens of post-vocalic /r/ were checked for the presence or absence of an audible /r/. Then, a similar check was made for the Arabic /r/'s in similar environments in order to detect whether the subjects have applied non-rhoticity onto their Arabic /r/'s. As for the variants that were other than a tap, trill, or an approximant, these revealed a wider repertoire for the bilinguals and will be discussed further in the results section.

Acoustic analysis was two-fold. First, all the /r/ tokens were inspected and identified as consisting of either formant-like features similar to the ones expected for approximants, or burst-like features that would identify the /r/'s as obstruents. Then a further acoustic analysis was undertaken of the approximant-like tokens and F1, F2, and F3 frequencies were measured in order to distinguish between [ɹ] and [ʋ]. It was expected that [ɹ] tokens would have lower F3 frequencies than [ʋ] tokens. A further acoustic analysis was also undertaken of the burst-like tokens in order to distinguish between trills (several bursts), taps (one burst), and weak taps (formant structure).

A total of 5229 tokens were analysed for this study, consisting of around 1000 /r/ tokens from the IViE corpus which were auditorily analysed, and 4229 /r/ tokens from this study which were initially auditorily analysed, and then around 1500 of these tokens were analysed acoustically. However, results from the acoustic analysis will not be presented quantitatively, but will rather be used as qualitative support to the auditory analysis.

In each of the sections that will follow group results will be presented first in order to show general trends, and then individual and detailed results will follow.

4.7.English results

4.7.1 Adults: group results

Results for the adults from the IViE corpus are discussed first, followed by results from the parents in this study in order to assess in more detail the specific targets available for both the monolingual and the bilingual children in the study. Starting with the data from the IViE corpus, Figure 4.1 shows the patterns found for the ten speakers from Leeds. As can be seen, the use of the approximant [ɹ] is categorical for all ten speakers and suggests that a change has taken place in the Leeds accent with respect to /r/ production, which has been described as a tap (Wells, 1982: 368). Note, however, that all ten speakers are 16 years old and their production does not necessarily reflect that of older generations. All ten speakers also had a non-rhotic accent.

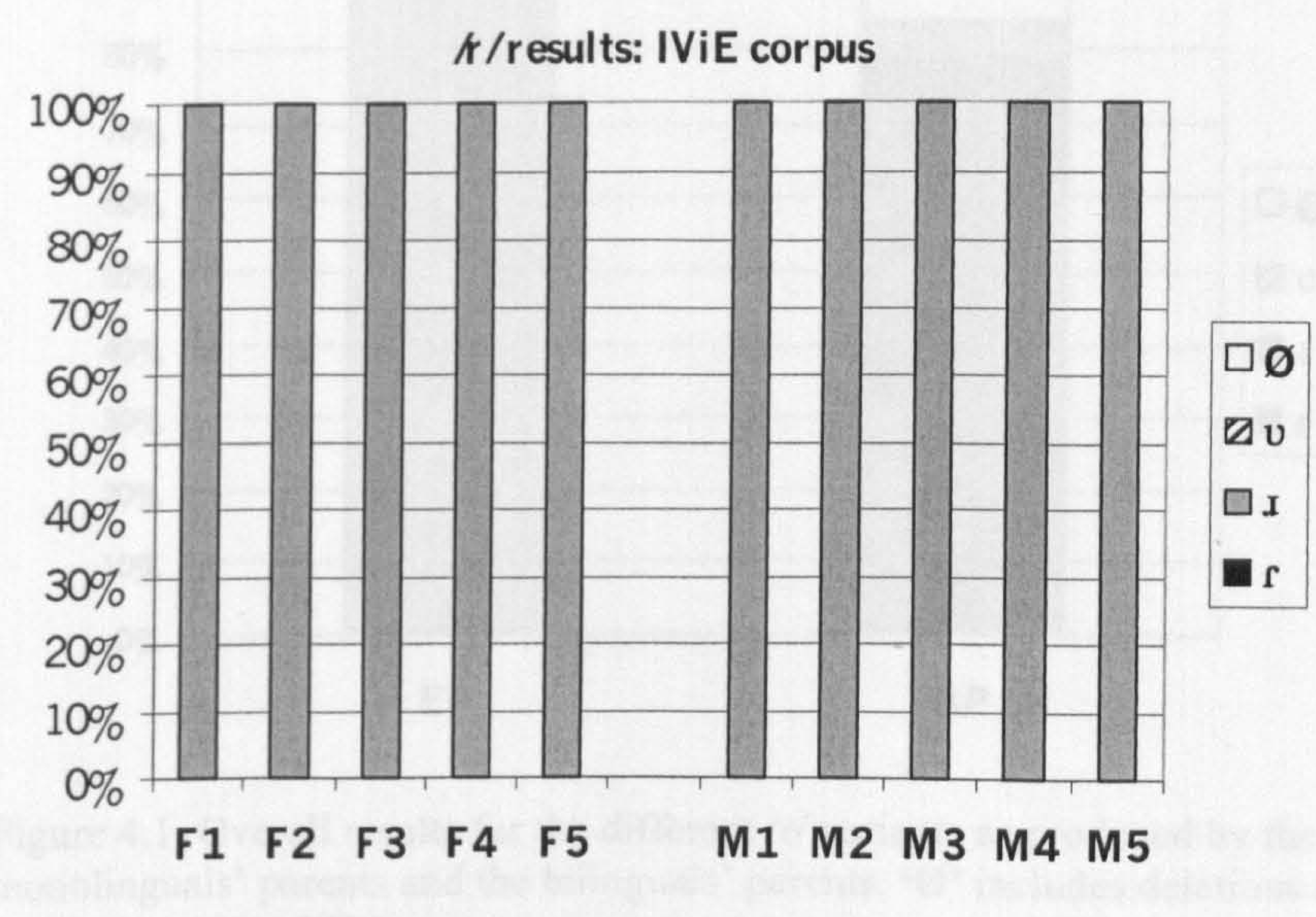


Figure 4.1: Results for the different /r/ variants as produced by the ten Leeds speakers from the IViE corpus. N = 1000.

Moving on to the parents in my study, Figure 4.2 shows group results for the /r/ variants used by the monolinguals' parents and the bilinguals' parents. Some categories of variants have been collapsed together in order to concentrate on the obstruent-*versus*-approximant pattern first. As can be seen, the overwhelming variant used by the monolinguals' parents is again the approximant [ɹ], and all six of them have a non-rhotic accent. There was a small percentage of the labial variant too, but this will be commented on when we look at the individual results in Figure 4.2. The bilinguals' parents, on the other hand, display the typical behaviour of L2 speakers whose L1 patterns interfere with their L2 by producing just under 80% of their /r/'s in English as taps or trills and by having a rhotic accent. Therefore the patterns shown in Figure 4.1 for the bilinguals' parents apply to /r/'s found in both pre- and post-vocalic positions The bilinguals' parents did produce a small number of [ɹ]s, along with a number of non-rhotic productions (hence the zero-realisations in post-vocalic environments). This indicates that they are aware of the English /r/ production patterns (which is not surprising when one considers that they have all been living in the UK for at least ten years), but do not or cannot produce them consistently due to the influence of Arabic /r/ patterns. What is important to note, though, is that the kind of variety displayed in Figure 4.1 constitutes part of the input that the bilingual children in this study are exposed to, with both the monolinguals' parents and their own parents being part of their surroundings.

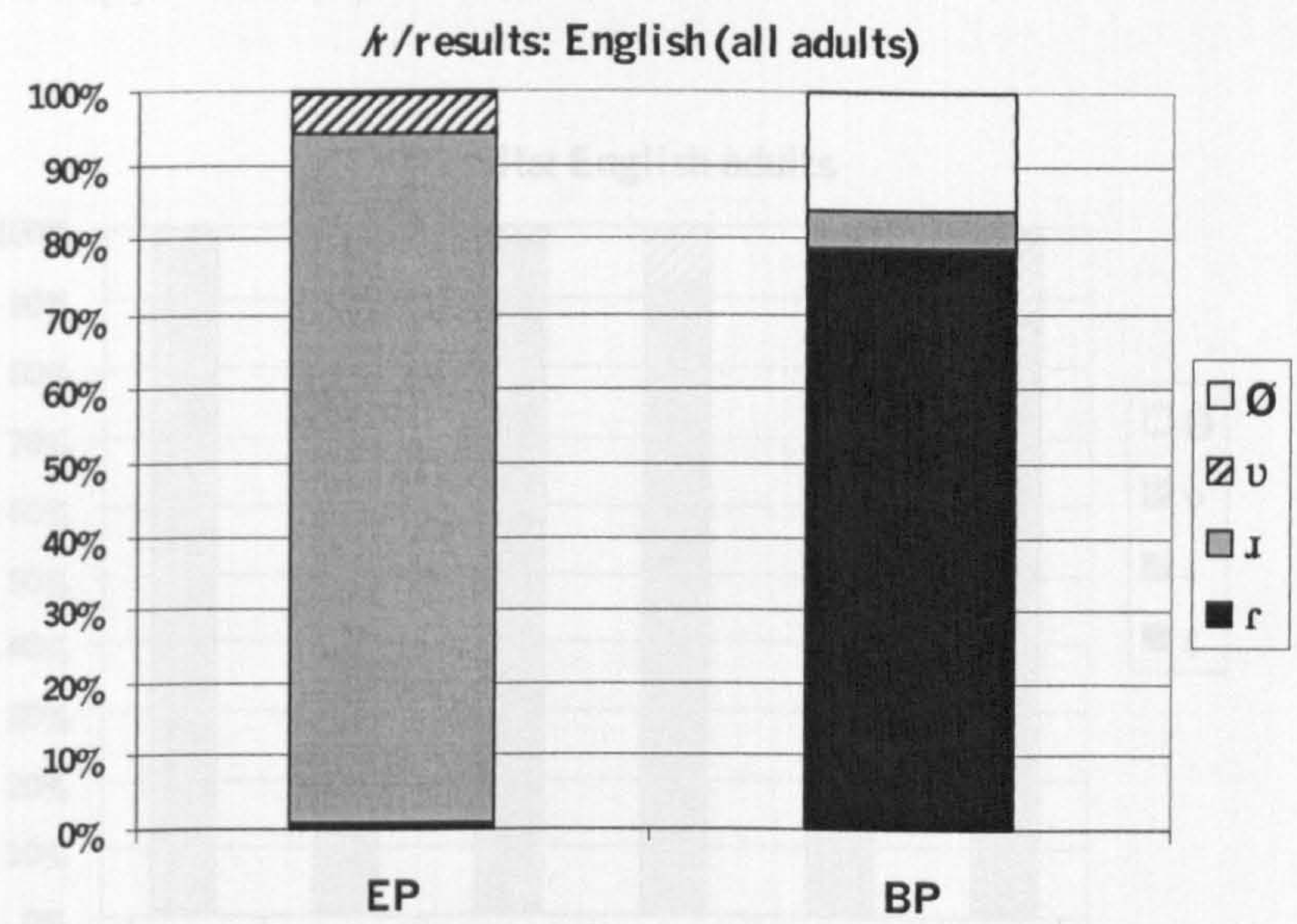


Figure 4.1: Overall results for the different /r/ variants as produced by the monolinguals' parents and the bilinguals' parents. 'Ø' includes deletions and other realisations. N = 1574.

4.7.2 Monolinguals' parents: individual results

As mentioned in the previous section, the monolinguals' parents mainly produced the [ɹ] variant (apart from two sporadic tokens by EF5 and EF7), which suggests that the approximant (and not the tap) is the variant used by this small Yorkshire community of speakers, although one has to be careful about generalizing the results since most of the parents are not originally from Yorkshire (see Chapter Two). Still, the fact that they have all been living in Leeds or York for approximately 10 years suggests that their speech patterns should be taken into consideration when one is describing the type of varieties available in the region in general, and in the bilingual subjects' environment in particular.

As for the small percentage of [ʋ]'s that were found in the monolingual parents' production, individual results (Figure 4.2 and Table 4.3) show that the labial variant was mainly produced by EM10, who originally comes from London where he grew up till he was 18. His use of [ʋ] is therefore not surprising, since the labial variant is a well-documented realization for /r/ in the South-East (Foulkes & Docherty, 2000; Hughes & Trudgill, 1996: 60). As for EF5, EM5, and EF10, the [ʋ]-like tokens that they produce were very small in number (Table 4.3), were mainly grouped under category 7 (weak [ʋ]), and mainly occurred as a result of an adjacent labial ('bread' [bʋɛd]; 'frog' [fʋɔg]) as noted elsewhere (Foulkes & Docherty, 2000). EM10, on the other hand, produced clearly identifiable [ʋ]'s (Table 4.3), and had additional realizations that were categorized as either weak [ɹ] or weak [ʋ].

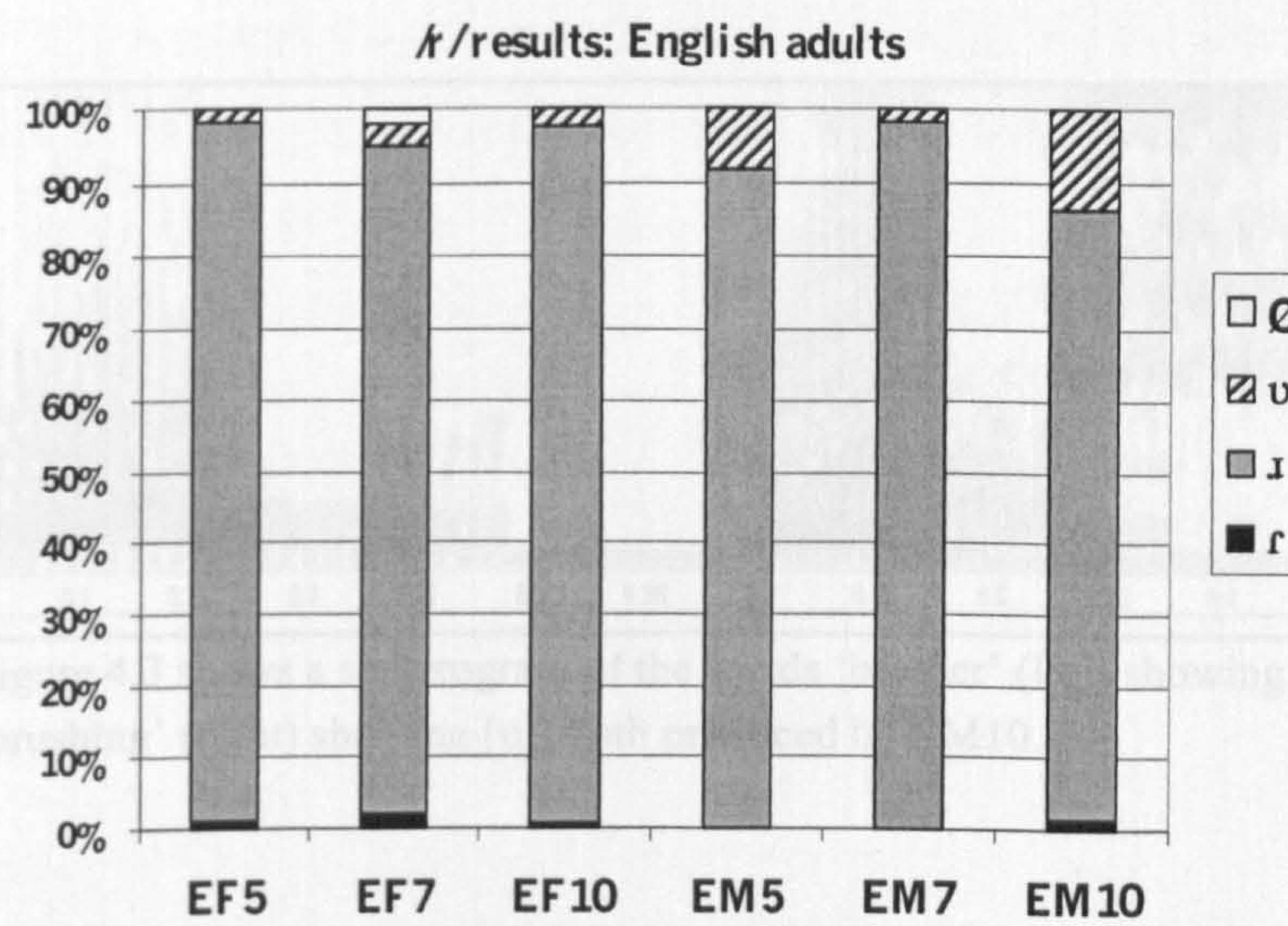


Figure 4.2: Results for the different /r/ variants produced by the monolinguals' parents. 'Ø' includes deletions and other realisations. N = 720.

Table 4.3: Detailed figures for the /r/ variants produced by the monolinguals' parents.

	EF5			EF7			EF10			EM5			EM7			EM10		
	read	story	N	read	story	N	read	story	N	read	story	N	read	story	N	read	story	N
r	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
ɾ	1	0	1	2	1	3	0	1	1	0	0	0	0	0	0	0	1	1
ɾ̥	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
ɹ	50	48	98	53	59	112	47	63	110	67	36	103	52	51	103	55	41	96
ɹ̥	0	5	5	2	11	13	0	1	1	4	3	7	0	9	9	11	3	14
weak ʋ	1	1	2	0	3	3	0	0	0	7	3	10	0	2	2	2	1	3
ʋ	0	0	0	0	1	1	3	0	3	0	0	0	0	0	0	13	2	15
Ø	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
other	0	0	0	2	1	3	0	0	0	0	0	0	0	0	0	0	0	0
Total	52	54	106	59	76	135	50	65	115	78	42	120	52	62	114	81	49	130

Acoustic analysis carried out on EM10's [ɹ] versus [ʋ] tokens reveals a tendency for F3 to have a higher value for the [ʋ] tokens, though more tokens in comparable contexts are needed to confirm this observation. The lack of F3 lowering in the labial variant has been reported elsewhere (Foulkes & Docherty, 1999; 2000), and suggests that [ʋ] lacks the strong retroflexion or tongue bunching normally typical of [ɹ]. Figure 4.3 shows a spectrogram of the words 'brother' and 'brushing' as produced by AM10, the first one with the alveolar approximant ([bɹʌðə]) and the second one with the labial one ([bʋʌʃɪŋ]). While in 'brother' F3 for [ɹ] is low and close to F2 (F3 = 11.56Z or 1627Hz), the [ʋ] in 'brushing' has clearly separate F2 and F3, with an F3 of 12.56Z or 1871Hz.

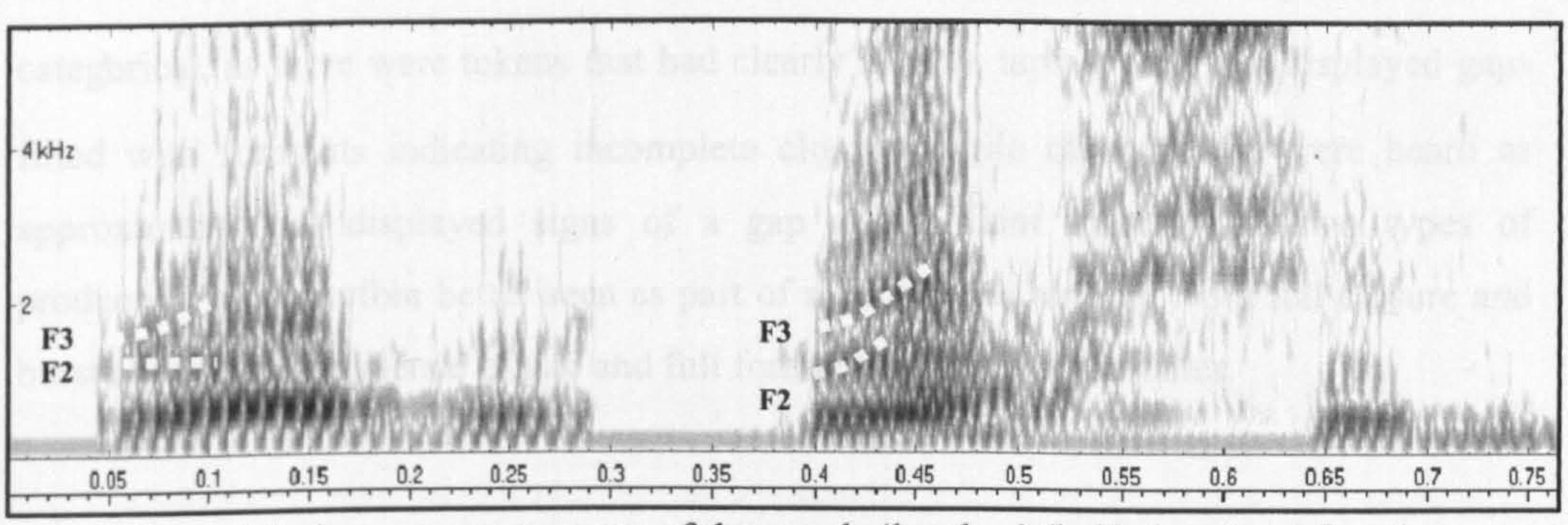


Figure 4.3 shows a spectrogram of the words 'brother' (left) showing [ɹ] and 'brushing' (right) showing [ʋ], both produced by EM10.

4.7.3 Bilinguals' parents: individual results

Not only do the bilinguals' parents mainly produce taps in their production of English /r/'s, but they also have post-vocalic productions in the majority of cases (Figure 4.4).

Since both taps and post-vocalic /r/'s are features of Arabic /r/'s, the behaviour of the bilinguals' parents can mainly be explained in terms of interference from their L1. The production of post-vocalic /r/'s may also reflect the type of English that the parents learned before moving to England.

There are, however, interesting observations with respect to sub-group and individual behaviour in the production of /r/'s (Table 4.4). First, three out of four of the L2 adults regularly produce a variant that sounds more like an approximant than a tap but that lacks the rounding and retroflexion that are typical of [ɹ]. For this reason, the variant has been labeled [ɾ̞] as it is assumed that it is produced following an incomplete or lack of contact that is typical of a tap articulation. Further acoustic investigation confirms that [ɾ̞] realizations are indeed more approximant in nature than stop-like due to the formant structure that they display in the majority of cases, but more importantly, that they are indeed different from the English [ɹ] due to the lack of F2 and F3 lowering that they exhibit (Figures 4.5 and 4.6). Figure 4.5 shows a spectrogram for the word 'cherries' produced with a typical tap [ɾ], showing a clear short gap and a burst, while Figure 4.6 shows a spectrogram for the word 'cherries' produced first by BM7 (left) with a weak tap [ɾ̞], and then by one of the monolinguals' parents, EM7 (right), with an approximant [ɹ]. Note the formant-like structure in medial position for both productions and low amplitude in the higher formants, but while there is no F2 and F3 lowering in BM7's production (F2 = 12.27Z or 1790Hz; F3 = 14.77 or 2605Hz), EM7's [ɹ] shows typical F2 and F3 lowering found for the bunched retroflex alveolar (F2 = 10.47Z or 1363 Hz; F3 = 12.64Z or 1892Hz). Of course, the difference between [ɾ] and [ɾ̞] was not always clear-cut or categorical, as there were tokens that had clearly audible taps [ɾ] but that displayed gaps filled with formants indicating incomplete closure, while other tokens were heard as approximants but displayed signs of a gap and/or faint bursts. The two types of production are therefore better seen as part of a continuum ranging from full closure and burst at one end to absence of gap and full formant structure at the other.

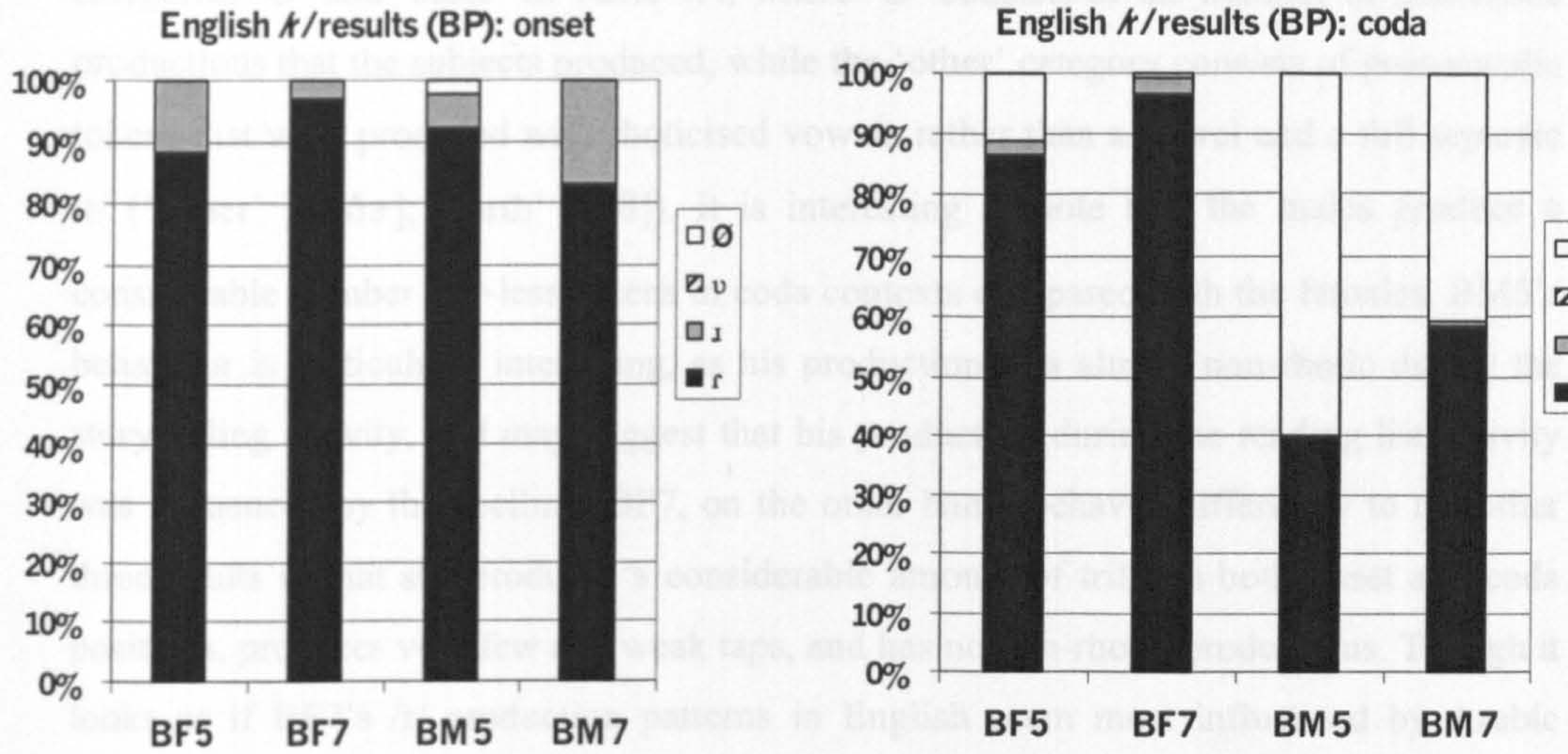


Figure 4.4: Results for the different /r/ variants produced by the bilinguals' parents. 'Ø' includes deletions and other realisations. N = 854.

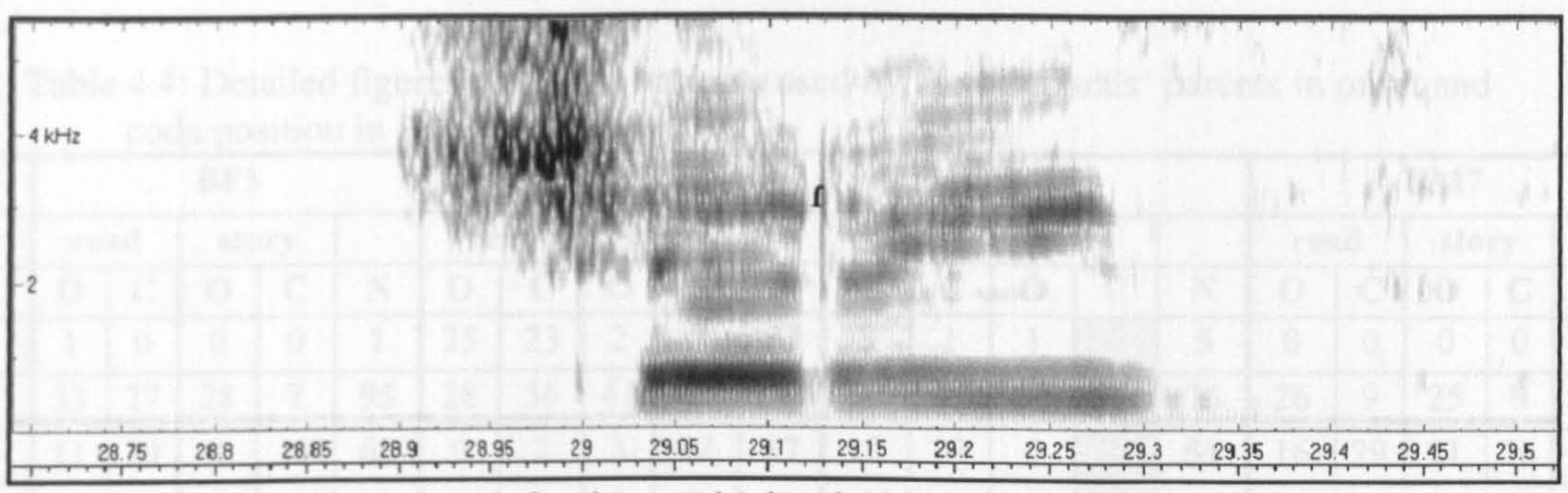


Figure 4.5: Spectrogram for the word 'cherries' produced with a tap [ɾ]

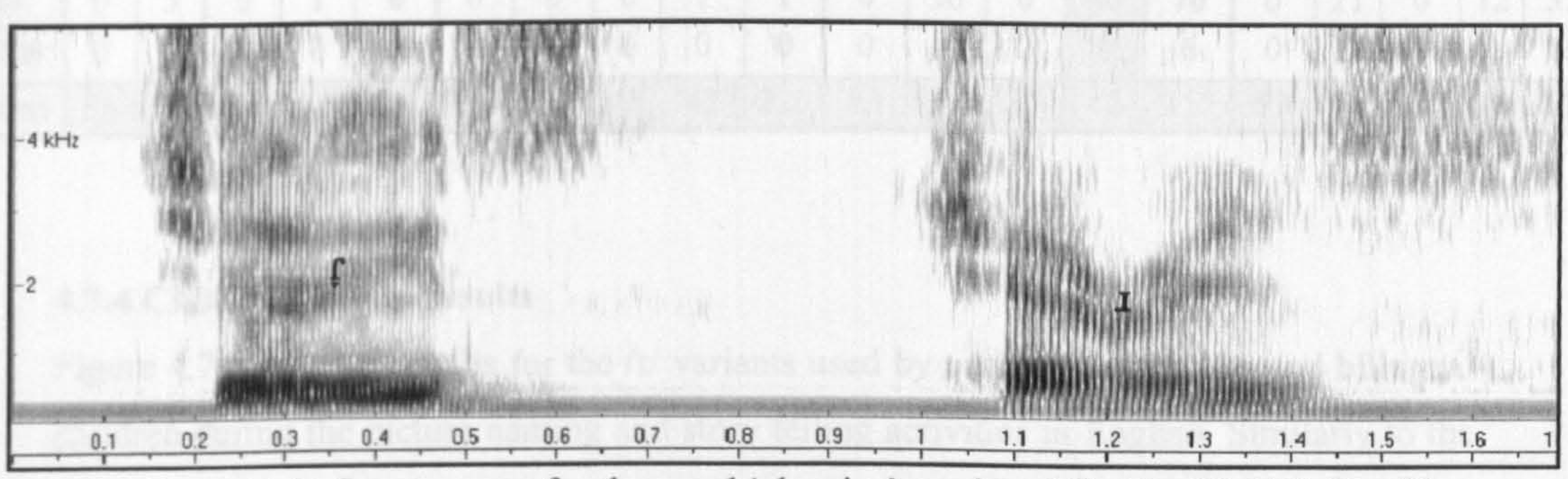


Figure 4.6: Spectrogram for the word 'cherries' produced first by BM7 (left) with a weak tap [ɾ] by EM7 (right) with an approximant [ɹ]

The weak tap [ɾ] was also found in the production of Arabic /r/ by the bilinguals' parents (see Section 4.8.1), but the fact it occurred more frequently in English and could be seen as an attempt on the part of the L2 adults to move their production closer to the English approximant. Further evidence to support this observation can be seen in the categories 'Ø' and 'other' in Table 4.4, where 'Ø' consists of the number of non-rhotic productions that the subjects produced, while the 'other' category consists of post-vocalic tokens that were produced with rhoticised vowels rather than a vowel and a full separate /r/ ('father' [fɑːðɚ]; 'earth' [ɝːθ]). It is interesting to note that the males produce a considerable number of r-less tokens in coda contexts compared with the females. BM5's behaviour is particularly interesting, as his production was almost non-rhotic during the story telling activity, and may suggest that his production during the reading list activity was influenced by the spelling. BF7, on the other hand, behaves differently to the other three adults in that she produces a considerable amount of trills in both onset and coda positions, produces very few any weak taps, and has no non-rhotic productions. Though it looks as if BF7's /r/ production patterns in English seem most influenced by Arabic compared with the rest of the L2 parents, strong trill and tap articulations are part of her idiosyncratic preference and she generally produces more trills in Arabic than any of the eight Arab-speaking adults in this study (see Section 4.8.1).

Table 4.4: Detailed figures for the /r/ variants used by the bilinguals' parents in onset and coda position in English.

	BF5					**BF7**					**BM5**					**BM7**				
	read		**story**			**read**		**story**			**read**		**story**			**read**		**story**		
	O	**C**	**O**	**C**	**N**	**O**	**C**	**O**	**C**	**N**	**O**	**C**	**O**	**C**	**N**	**O**	**C**	**O**	**C**	**N**
r	1	0	0	0	**1**	25	23	2	3	**53**	3	1	1	0	**5**	0	0	0	0	**0**
ɾ	33	27	28	7	**95**	28	56	41	24	**149**	30	17	28	1	**76**	26	9	25	9	**69**
ɾ̥	11	39	4	11	**65**	0	2	3	12	**17**	15	29	6	1	**51**	16	29	21	26	**92**
ɹ	10	2	0	0	**12**	1	1	2	2	**6**	5	0	0	0	**5**	11	0	7	1	**19**
Ø	0	5	0	1	**6**	0	0	0	1	**1**	0	30	0	40	**70**	0	21	0	12	**33**
other	0	5	0	0	**5**	0	0	0	0	**0**	0	4	2	0	**6**	0	18	0	0	**18**
Total	**55**	**78**	**32**	**19**	**184**	**54**	**82**	**48**	**42**	**226**	**53**	**81**	**37**	**42**	**213**	**53**	**77**	**53**	**48**	**231**

4.7.4 Children: group results

Figure 4.7 shows the results for the /r/ variants used by each of the English and bilingual children during the picture naming and story telling activities in English. Similarly to the presentation of the results for the adults, some categories of variants have been collapsed together in order to concentrate on the obstruent-*versus*-approximant pattern first. As can be seen, the overwhelming variant used by both the child groups is the approximant [ɹ], which shows that the bilinguals have not adopted the /r/ patterns that were produced by their parents and are allowing very little interference from Arabic (note the small percentage of taps produced). More interestingly, all three bilingual children have a non-rhotic accent in English, though all their parents are predominantly rhotic. Both groups of children also produced labial variants of /r/ and other realizations, but these will be discussed in more detail in the individual results.

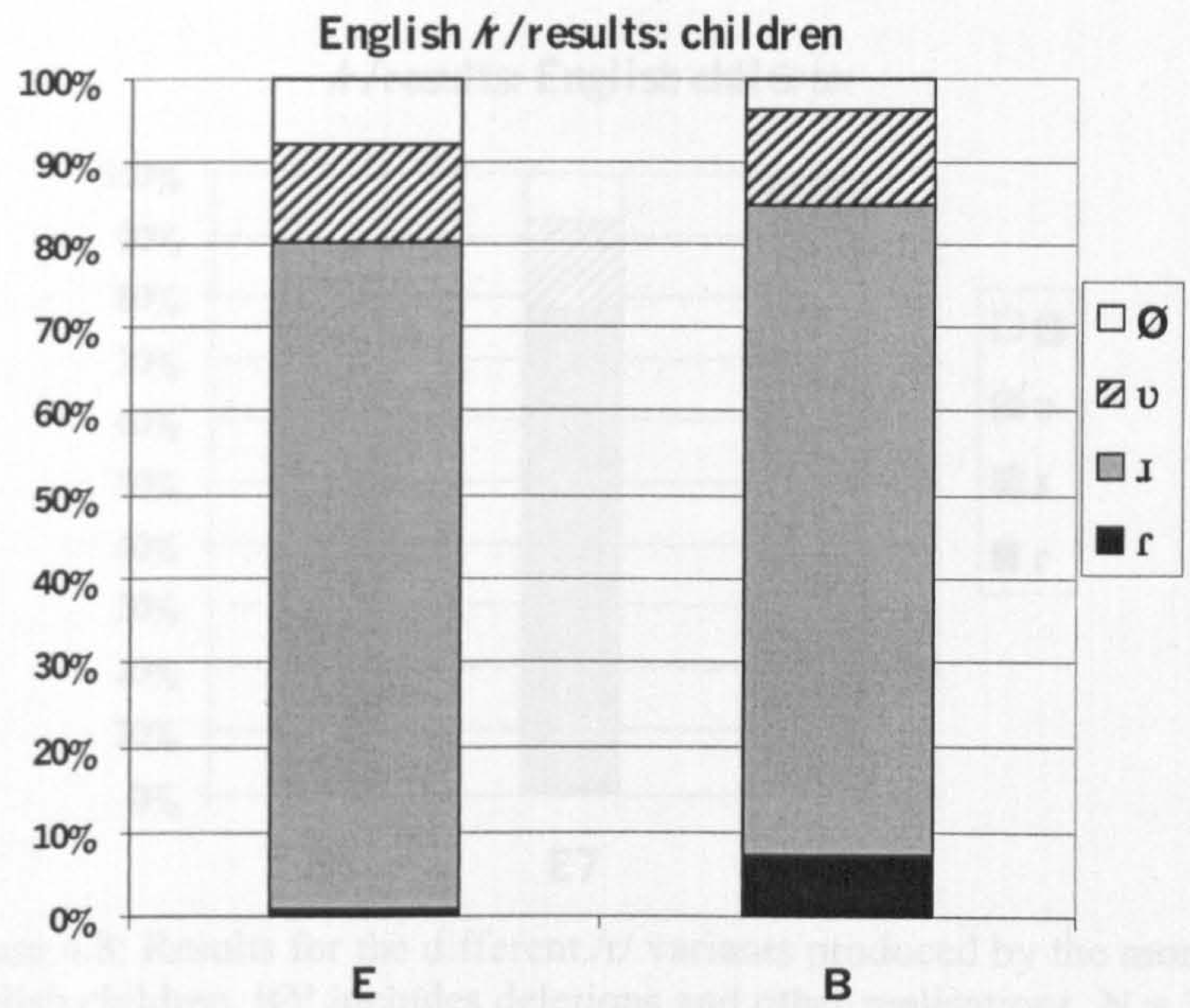


Figure 4.7: Results for the different /r/ variants produced by the English and bilingual children. 'Ø' includes deletions and other realisations. N = 430.

4.7.5 Monolingual English children: individual results

Figure 4.8 and Table 4.5 show the individual results for the monolingual children's /r/ production patterns. Apart from the overwhelming use of the [ɹ] variant by all the children, there are two main developmental patterns that seem to be taking place in their production as they grow older. First, the percentage of the labial variant [ʋ] gradually decreases as the age of the children increases. Such a result is to be expected knowing that the production of [ɹ] involves physically complex articulations and usually emerges late in children's speech. As mentioned in Section 4.1.4, children acquiring [ɹ] frequently replace it by [w] and [ʋ], and the children in this study are no exception[7]. The second developmental feature is /r/ deletion in onsets which, like [ʋ] production, decreases as the age of the children increases. Deletion is also expected among children acquiring /r/, and normally takes place in consonant clusters and in medial and final positions (in rhotic

[7] Although [ʋ] was also found in the production of one of the adults, there is no evidence in the literature or from the other adults in this study to show that the labial variant is part of the accent of the community.

accents). Examples from this study include [g̊iːn] 'green', [ˈpaɪde] 'Friday', [ˈdʒanui] 'January', and [fɔg] 'frog' by E5 (note other developmental features like devoicing of /g/ in 'green' and stopping of /f/ in 'Friday'), and [ˈbəʊkən] 'broken', [ˈeɪpəɫ] 'April' by E7.

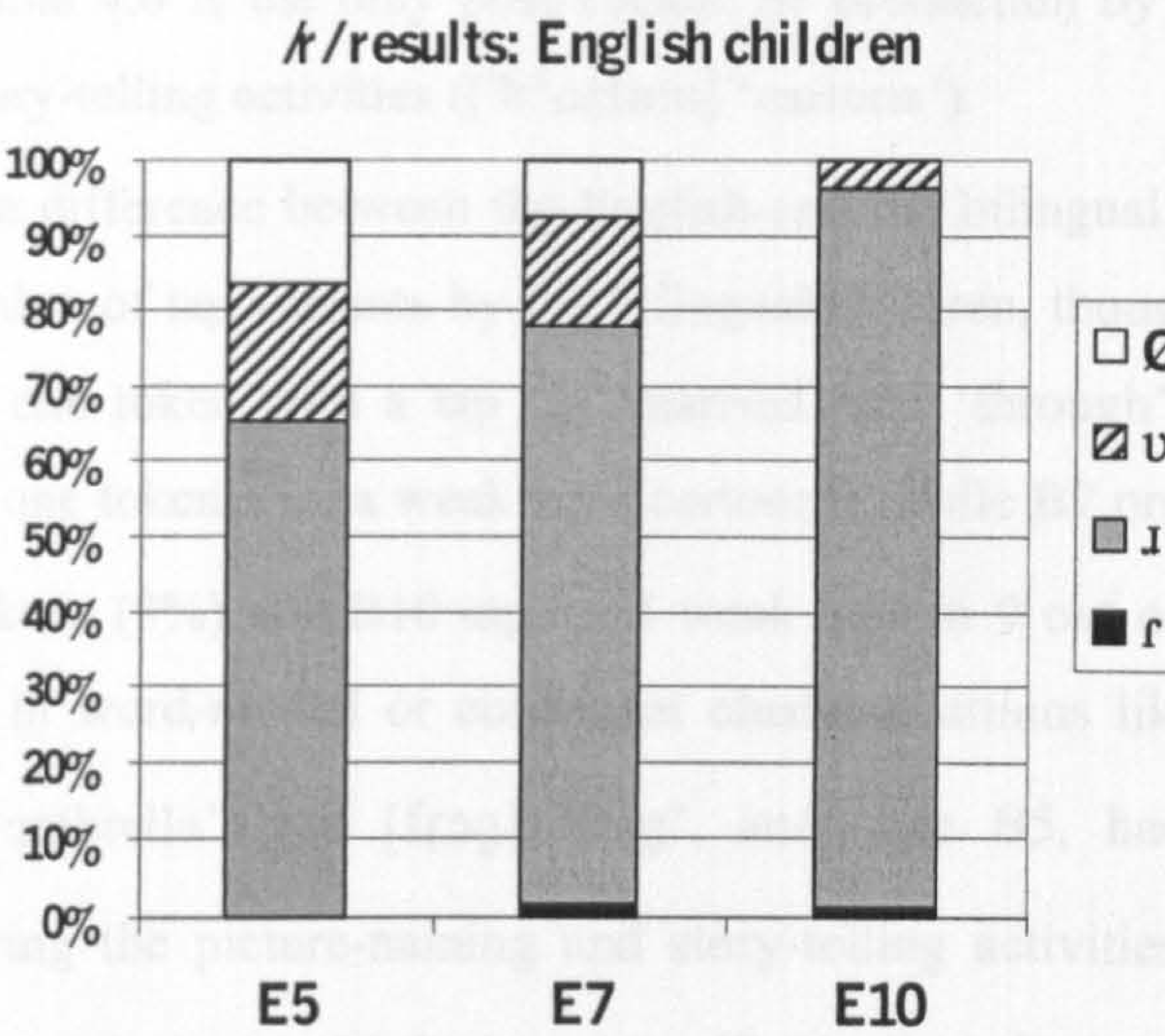


Figure 4.8: Results for the different /r/ variants produced by the monolingual English children. 'Ø' includes deletions and other realisations. N = 213.

Table 4.5: Detailed figures for the /r/ variants used by the monolingual English children during the picture-naming (pic) and story-telling (story) activities.

	E5			E7			E10		
	pic	story	N	pic	story	N	pic	story	N
ɾ	0	0	**0**	1	0	**1**	0	1	**1**
ɾ̥	0	0	**0**	0	0	**0**	0	0	**0**
ɹ	36	11	**47**	33	15	**48**	48	21	**69**
ɹ̥	0	0	**0**	1	0	**1**	0	4	**4**
weak ʋ	4	1	**5**	5	0	**5**	1	0	**1**
ʋ	5	3	**8**	2	2	**4**	1	1	**2**
Ø	6	6	**12**	3	1	**4**	0	0	**0**
other	0	0	**0**	1	0	**1**	0	0	**0**
Total	**51**	**21**	**72**	**46**	**18**	**64**	**50**	**27**	**77**

4.7.6 Bilingual children: individual results, English produced in the English-only sessions

Figure 4.9 and Table 4.6 show the individual results for the bilingual children's /r/ production patterns when speaking English. At first sight, the patterns for the bilinguals seem very similar to those of the monolingual English children, mainly with regards to the use of [ɹ] and the gradual decrease in the use of [ʋ] and deletions as the children grow older. It is interesting to note that the youngest bilingual (B5) produced fewer deletions than the English 5-year-old (3% for B5 (2 out of 63 tokens) compared with 17% for E5 (12 out of 72 tokens)). B5's 2 tokens listed in the category 'other' (Table 4.9) consist of [w] realizations in ['mɪwə] 'mirror' and ['skwi:mɪŋ] 'screaming'. One token that was not included in Table 4.6 is the only post-vocalic /r/ production by B5 during the picture-naming and story-telling activities (['kʰɑɾ̥tu:n] 'cartoon').

The main difference between the English and the bilingual child groups is the use of a small number of tap variants by the bilingual children, though each of E7 and E10 also produced one token with a tap (in 'married' and 'through' respectively). B5 also produced only one token with a weak tap ('cartoon'), while B7 produced weak taps [ɾ̥] in 6 out of 77 tokens (8%) and B10 taps and weak taps in 9 out of 77 (12%). B7 mainly produced taps in word-medial or consonant cluster positions like in ['kʰaɾ̥ət] 'carrot', [əm'bɾ̥ɛlla][8] 'umbrella', and [fɾ̥ɔg̊] 'frog', and, like B5, had one post-vocalic /r/ production during the picture-naming and story-telling activities (['ɔtʰɚ] 'otter'). B10 produced taps mainly in word-initial position like in ['ɾʉstʰəɾ̥] 'rooster' and [ɾɛd] 'red', but also in consonant clusters like [θɾʉ] 'through', and [əm'bɾ̥ɛla] 'umbrella'. B10 also had two post-vocalic /r/ tokens during the picture-naming and story-telling activities, ['ɾʉstʰəɾ̥] 'rooster' and [dʒɑ:ɾ̥] 'jar'.

Still, apart from the few tap productions which constitute only a small percentage of the bilinguals' overall /r/ production in English, the three bilingual children do have overall similar /r/ patterns to those of the monolingual English children in this study. But that is not the whole picture (see Section 4.7.8).

---

[8] The gemination of /l/ is similar to how B7's mother (BF7) produces the word 'umbrella' and may be an influence from Arabic, where /l/ is frequently geminated.

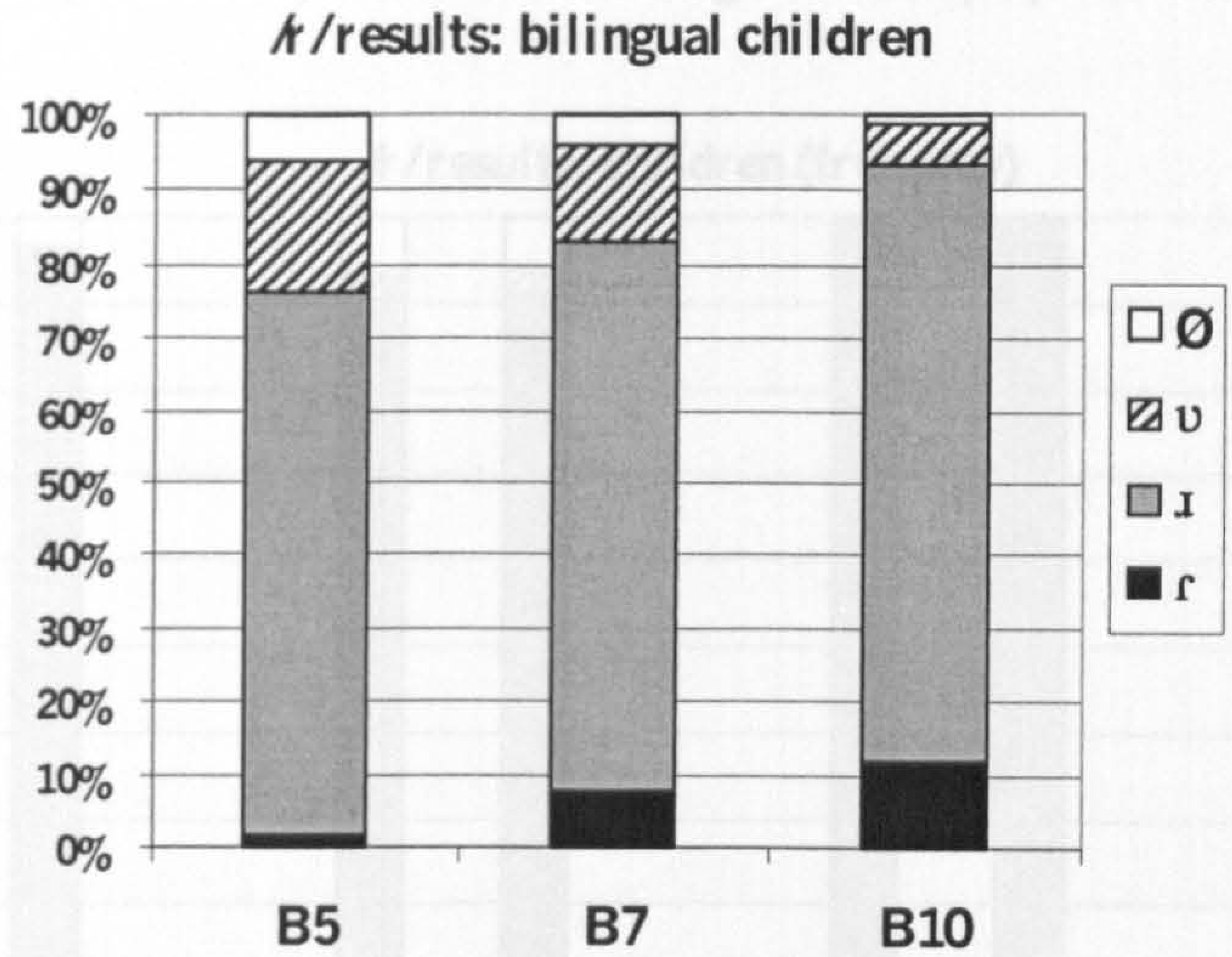


Figure 4.9: Results for the English /r/ variants produced by the bilingual children during the English sessions. 'Ø' includes deletions and other realisations. N = 217.

Table 4.6: Detailed figures for the /r/ variants used by the bilingual children during the picture-naming (pic) and story-telling (story) activities in English.

	B5			B7			B10		
	pic	story	N	pic	story	N	pic	story	N
ɾ	0	0	**0**	0	0	**0**	5	2	**7**
ɾ̥	0	1	**1**	2	4	**6**	1	1	**2**
ɹ	25	22	**47**	33	23	**56**	40	23	**63**
ɹ̥	0	0	**0**	0	2	**2**	0	0	**0**
weak ʋ	3	1	**4**	8	0	**8**	3	0	**3**
ʋ	4	3	**7**	2	0	**2**	1	0	**1**
Ø	2	0	**2**	1	1	**2**	0	0	**0**
other	2	0	**2**	0	1	**1**	1	0	**1**
Total	**36**	**27**	**63**	**46**	**31**	**77**	**51**	**26**	**77**

4.7.7 Children: free-play sessions

Since I conducted the picture-naming story telling activities, material from the free-play sessions between the children was used in order to support the findings reported in Section 4.7.6. As mentioned in Chapter Two, each of the bilingual children was recorded playing with a monolingual friend of the same age and B7 and B10 were also recorded playing together in order to test any possible difference in the bilinguals' linguistic behaviour depending on whether they are interacting with monolinguals or bilinguals. Figure 4.10 and Table 4.7 show the /r/ patterns produced by each of the 6 children during the free play sessions whereby each bilingual was paired with a monolingual, along with the /r/ patterns produced by B7 and B10 during their free play session.

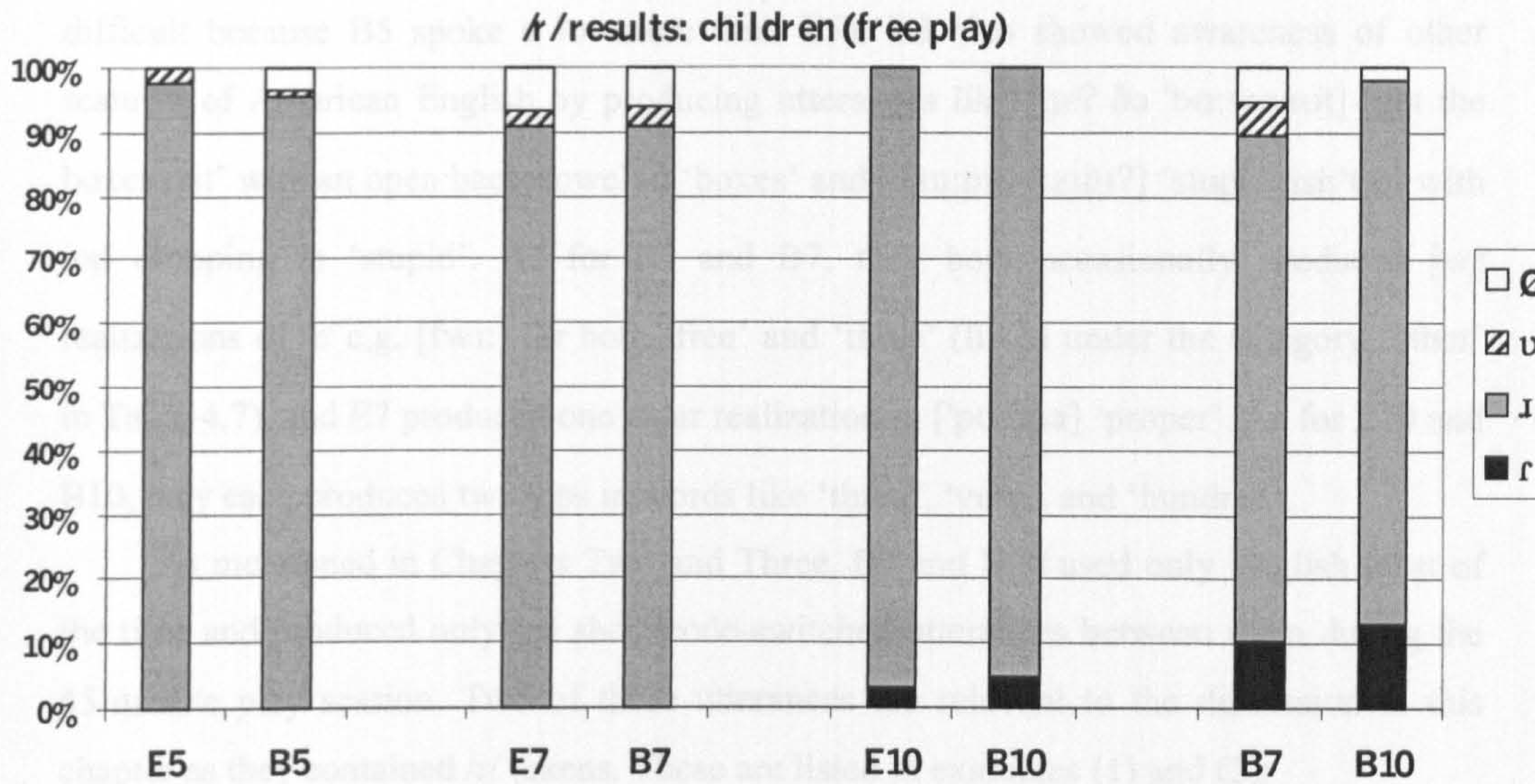


Figure 4.10: Results for the /r/ patterns that were found during the paired free-play sessions between the monolingual and the bilingual children, and during the free-play session between B7 and B10 (far right). N = 391.

Table 4.7: Detailed results for /r/ pattern during the paired free-play sessions between the children.

	N	ɾ	ɽ	ɹ	ɻ	weak ʋ	ʋ	Ø	other
E5	43			39	3		1		
B5	93			87	2	1		3	
E7	44			40			1		3
B7	33			28	2	1		1	1
E10	67	2		65					
B10	44	2		42					
B7	19	1	1	14	1	1		1	
B10	48	5	1	41					1

The patterns that emerged from the free-play sessions support the results that were found in Section 4.7.5 and 4.7.6 and that show that the approximant [ɹ] is the majority variant for both monolingual and bilingual children. Interesting observations include the fact that E5 and B5 produced a number of post-vocalic /r/'s during their role-play sessions with their dolls while they were imitating the American accents of cartoon characters they had just been watching before the beginning of the recording session.

Recall that E5's father, EM5, noted that his daughter often came up with American expression although he did not know where she learned them. B5 produced many more American-sounding utterances with rhoticised vowels than E5 (though comparison is difficult because B5 spoke a lot more than E5). B5 also showed awareness of other features of American English by producing utterances like [gɛʔ ðə ˈbɑːsəzaʊt] 'get the boxes out' with an open back vowel in 'boxes' and [ˈstuːpɪd ˈɪzɪnɪʔ] 'stupid, isn't it' with jod dropping in 'stupid'. As for E7 and B7, they both occasionally produced [w] realizations of /r/ e.g. [fwiː] for both 'free' and 'three' (listed under the category 'other' in Table 4.7), and E7 produced one velar realization in [ˈpɰɔpə] 'proper'. As for E10 and B10, they each produced two taps in words like 'three', 'very', and 'hundred'.

As mentioned in Chapters Two and Three, B7 and B10 used only English most of the time and produced only ten short code-switched utterances between them during the 45-minute play session. Two of these utterances are relevant to the discussion in this chapter as they contained /r/ tokens. These are listed in examples (1) and (2):

(1) B7: *Cause of you,* ħmaːɾ
Cause of you, donkey.

(2) B10: *Which colour* [kʰʌləɾ] ʔane
Which colour me?
Which colour is mine?

In (1), B7 produced the expected /r/ variant for the word 'donkey' in Arabic, while B10 produced a post-vocalic tap in the word 'colour' in English. B10 actually produced four other tokens in English-only utterances with post-vocalic /r/'s, though his production was non-rhotic during the session with E10 and during the English-only sessions with me. Still, apart from the small amount of post-vocalic productions by B10 and the few tap productions by both B7 and B10 (Table 4.7), it could be said that, on the whole, their productions patterns when playing together are still largely similar to those they exhibit when playing with their English friends.

4.7.8 Bilingual children: individual results, English produced during the Arabic sessions

As mentioned in Chapter Three, all the English tokens from the Arabic sessions were analysed separately due to the stark differences in the patterns that they display compared with English spoken in the English sessions. Examples (3) and (4) illustrate the types of code-switches that occurred.

e.g. (3) Mother (pointing at a dress): [ʃu haɪda]?
What that (masc.)?
What is that?

Child: [dɾɛs]
dress

e.g. (4) Child (describing a an action): [ˈnɑtˤtˤɪt mɪnɪl dʒɑɾ]
*jump-past-fem. out-of-the **jar***
she jumped out of the jar

Figure 4.11 shows the individual results for the bilingual children's /r/ production patterns in English during the Arabic sessions. There is a stark contrast between the results in this figure and those discussed in Figures 4.9 and 4.10, as the main variant used by the bilinguals this time is the tap rather than the approximant, and are a lot of post-vocalic productions (Tables 4.8 and 4.9). The two patterns of production in Figures 4.9 and 4.11 can only be seen as belonging to two different language modes for the bilinguals, and this issue will be discussed further in Chapter Six.

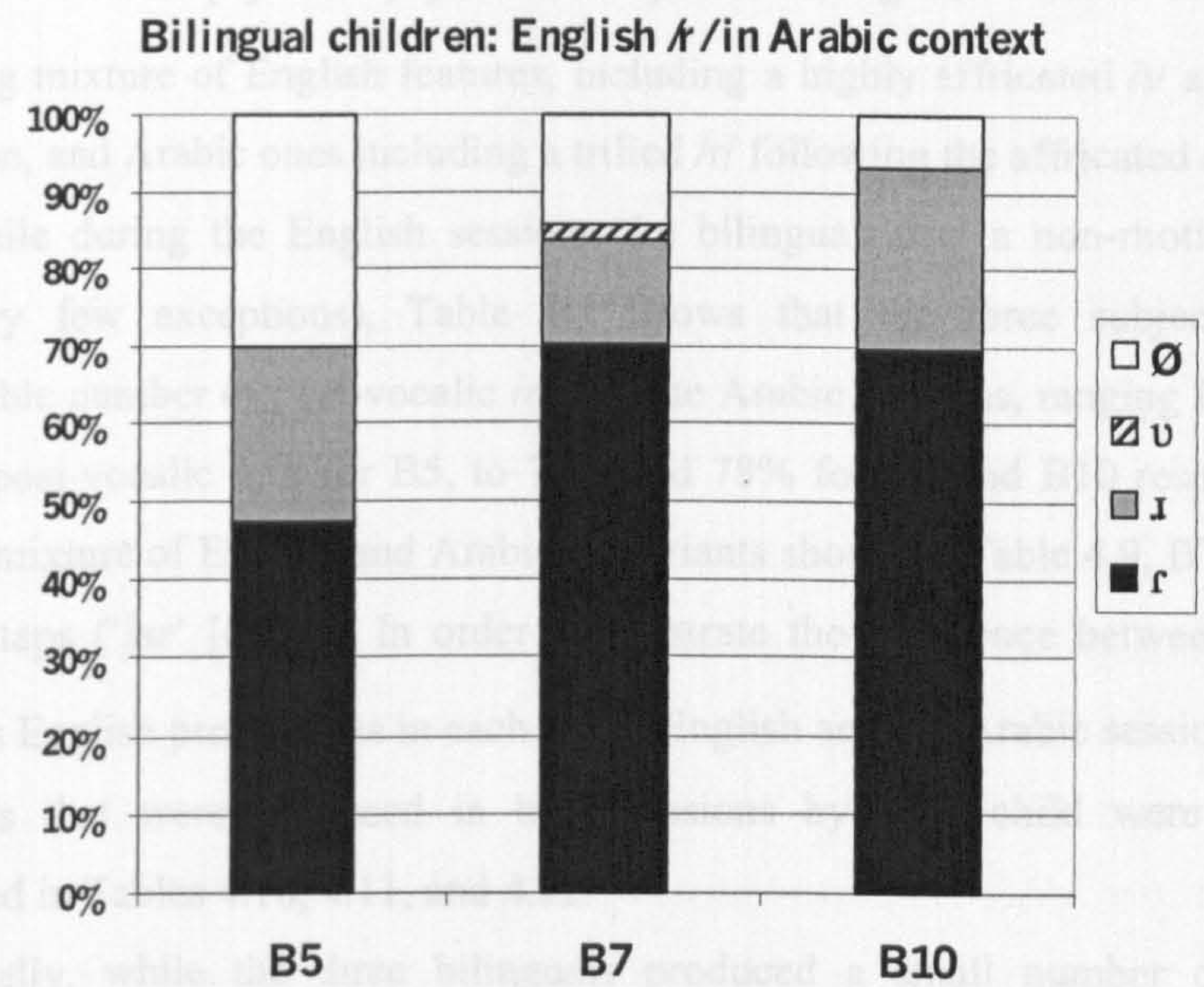


Figure 4.11: Results for the English /r/ variants produced by the bilingual children during the Arabic sessions. 'Ø' includes deletions and other realisations. N = 180.

Table 4.8: Detailed figures for the /r/ variants used by the bilinguals in onset positions.

Onset	N	r	ɾ	ɽ	ɹ	ʋ	Ø	other
B5	25	3	7	3	8	0	1	3
B7	47	0	24	10	11	2	0	0
B10	26	0	11	4	10	0	0	1
Total	**98**	**3**	**42**	**17**	**29**	**2**	**1**	**4**

Table 4.9: Detailed figures for the /r/ variants used by the bilinguals in coda positions.

Coda	N	r	ɾ	ɽ	ɹ	ʋ	Ø	other
B5	16	0	5	1	1	0	9	0
B7	49	0	20	13	2	0	12	2
B10	17	2	9	4	0	0	2	0
Total	**82**	**2**	**34**	**18**	**3**	**0**	**23**	**2**

As can be seen from Table 4.8 the bilinguals are making use of both the Arabic and the English /r/ categories during the Arabic sessions. In terms of other realisations, B5's production displays developmental features that are normally typical of (i) English acquisition, e.g. affrication ([tʃamp] 'tramp'), (ii) Arabic acquisition, e.g. lateralisation (['sᵊkli:mɪŋ] 'screaming') or (iii) both, e.g. /r/ deletion (['tʃɛi:h] 'cherry'). It is interesting to note that while there was no /r/ lateralisation in B5's production during the English-only sessions, she did produce it in the Arabic ones. On the other hand, while B5 did not produce any trills when speaking Arabic (see Section 4.8.2), she did produce three English tokens with trills during the Arabic sessions. Figure 4.12 shows a spectrogram of the word 'trainer' ['tʃreɪnəʰ] produced by B5 during the Arabic session, with an interesting mixture of English features, including a highly affricated /t/ and a non-rhotic production, and Arabic ones including a trilled /r/ following the affricated /t/.

While during the English sessions the bilinguals had a non-rhotic accent (apart from very few exceptions), Table 4.9 shows that the three subjects produced a considerable number of post-vocalic /r/'s in the Arabic sessions, ranging from 44% of all possible post-vocalic /r/'s for B5, to 72% and 78% for B7 and B10 respectively. Apart from the mixture of English and Arabic /r/ variants shown in Table 4.9, B7 also produced retroflex taps ('jar' [dʒɑɽ]). In order to illustrate the difference between the bilingual children's English productions in each of the English and the Arabic sessions, a sample of the words that were produced in both sessions by each child were extracted and transcribed in Tables 4.10, 4.11, and 4.12.

Finally, while the three bilinguals produced a small number of the alveolar approximant typical of their production in the English-only sessions, B7 is the only child who also produced two tokens with a labial approximant, while B10 produced one token with a retroflex approximant ([ɖɻɛs] 'dress'), which adds to the variety of /r/'s produced by the bilinguals in this Arabic sessions.

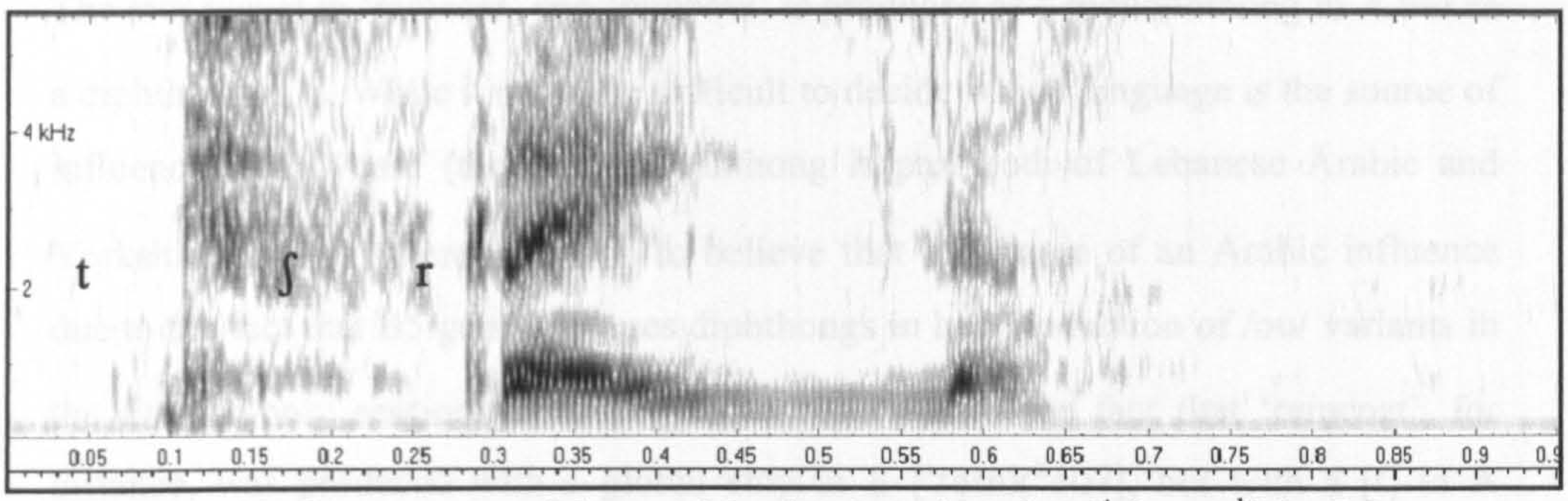


Figure 4.12: Spectrogram of the word 'trainer' produced as [ˈtʃreɪnəʰ] by B5 during the Arabic session.

Table 4.10: English tokens produced by the B5 in Arabic sessions compared with different productions of the same tokens in English sessions

		A	B
B5	**Gloss**	**Produced in Arabic sessions**	**Produced in English sessions**
Pre-V	raspberry	ˈɾa̤ːz̥b̥ᵊɾiː	ˈɹaz̥b̥ᵊiː
	raincoat	ˈɾe̦ɪnkʰot	ˈɹe̦ɪnkʰəʊʔ
	rainbow	ˈəɾeːnːboː	ˈɹeɪnbəʊ
Post-V	shark	ʃɑːɹk	ʃɑːk
	pepper	ˈpʰɛpʰəɾ	ˈpʰɛpə
	jumper	ˈdʒʌmpʰəɾ	ˈdʒʊmpʰəh
	scarf	skɑɾf	skɑːf
	garden	ˈɡɑːɾtən	ˈɡɑːdən / ˈɡɑːd̥ən
	thunder	ˈθʌndəɾ	ˈθʊnda
	marbles	ˈmaɾbəz̥	ˈmɑːbɫ̩z̥

As can be seen from Table 4.10, while the tokens in the English sessions consist solely of English sound features (taking into consideration developmental ones), the tokens produced during the Arabic sessions display a mixture of features from both English and Arabic sound patterns, affecting both /r/'s and neighbouring sounds in the relevant words. In order to avoid repeating the expressions 'Arabic session' and 'English session', the former is henceforth referred to as 'A' and the latter as 'E'. Following are some of the observations made about the patterns that differ in their production from one session to the other.

- The /a/ vowel in 'raspberries' produced in A is raised and sounds closer to the quality of an Arabic [a̤ː] produced for e.g. in proper names ([ˈma̤ːzən] 'Mazen').
- While the /l/ in 'marbles' is dark and syllabic when produced in E, it is replaced with a schwa in A, which could be seen as a strategy to break /l/ syllabicity into /ə + l/, except that the /l/ has been deleted as well. Schwa insertion before syllabic liquids and nasals is common in L2 speech (including that of the bilinguals' parents).

- The /əʊ/ vowel in 'raincoat' and 'rainbow' is produced as a monophthong in A but as a diphthong in E. While it might be difficult to decide which language is the source of influence in this case (the [o] monophthong is part both of Lebanese Arabic and Yorkshire English) there is reason to believe that it is more of an Arabic influence due to the fact that B5 generally uses diphthongs in her production of /əʊ/ variants in the English-only sessions. More evidence comes from the fact that 'raincoat', for instance, was produced with a glottal stop in E ['ɹe̹ɪnkʰəʊʔ], but with a [t] in A ['ɾe̹ɪnkʰot], making the latter realisation altogether more Arab-sounding than English. It is interesting to note how English-accent features like glottalling in 'raincoat' and [ʊ] production in 'jumper' and 'thunder' are only produced during the English sessions.
- VOT changes will be discussed in Chapter Five.

Table 4.11: English tokens produced by B7 in Arabic sessions compared with different productions of the same tokens in English sessions

		A	E
B7	Gloss	Produced in Arabic sessions	Produced in English sessions
Pre-V	red	ɾɛd̥	ɹɛd̥
	grandma's	gᵊ'ɾãmɑz̥	'gɹanniz̥
	cherries	'tʃeɾi:	'tʃeiz̥
	grapes	gɾeɪps	gɹe·ps
	carrot	'kɐɾ̥ət	'kʰaɹət
	fridge	fᵊ'ɾɪdʒ	'fɹ̥ɪdʒ
	drumming	d:ə'ɾəmɪŋ	'd^{3}ɹʊmɪŋg
Post-V	butterfly	bətəɾ̥'flaɪ	'bʌtˀflaɪ
	cucumber	ku'kʌmbɚ	'kʰʉkˣʌmbə
	beer	b:iəɾ	bɪə
	butter	ɓʌ'tʰəɾ	'ɓʌtʰə
	fingers	'fiŋgəɾz̥	'fiŋgəz̥
	circus	'sɪɾkəs	'sɜ:kəs
	earth	ɜɾθ	ɜ:θ
	purple	'pɜ:pəl	'pɜ:pɵ
	singer	'sɪŋgaɾ	'sɪŋgə
	waiter	'weɪtɑ̱ɾ̥	'we̹ɪtʰə
	marble(s)	'ma:lbəɾ	'mɑ:bɫz̥

Similarly to B5, the tokens produced during the Arabic sessions by B7 display a mixture of features from both English and Arabic sound patterns, affecting both /r/'s and other sounds in the tokens analysed. Apart from the dominance of taps and postvocalic

/r/'s in the productions from A, there are noticeable differences between the productions of vowels, consonants, consonants clusters and stress patterns in each language mode.

- In terms of vowel changes, the first vowel in 'carrot' is produced with a central quality [ɐ] in A and a front one typical of English /a/ in E (Figure 4.13); [ɐ] is a typical realisation of /a/ in Arabic.
- Similarly, the /uː/ vowel in 'cucumber' is realised with a back quality typical of Arabic [u] in A, while it has a more central quality typical of English English production of /u/ in E (note yod-dropping in both productions).

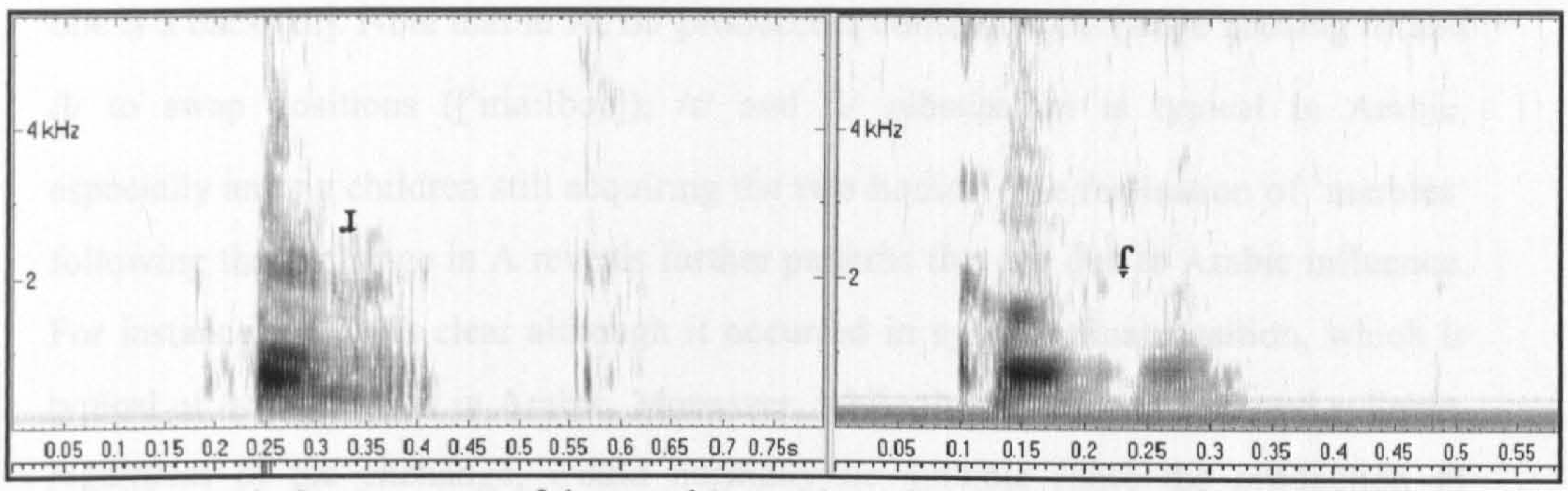


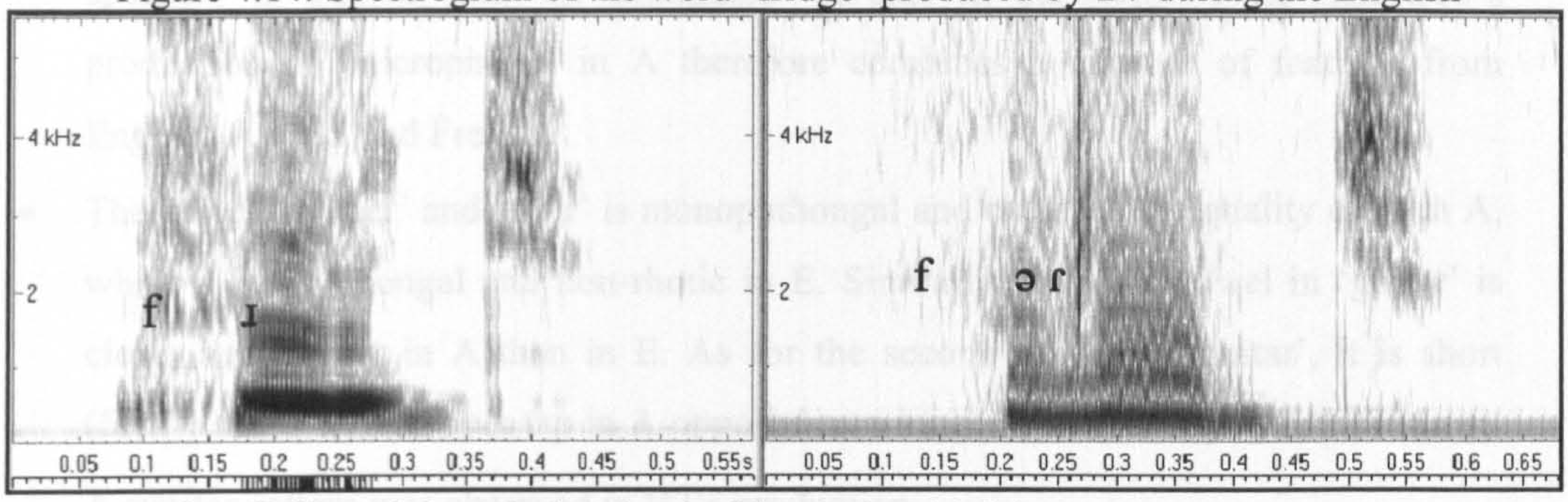
Figure 4.13: Spectrogram of the word 'carrot' produced by B7 during the English session (left), and then during the Arabic session (right).

- 'beer' is a loan word in Arabic normally produced as [biːɾa]. In B7's production in A, the vowel is a diphthong and sounds closer to an English production but with a high front [i] rather than [ɪ], therefore combining features from both languages, especially that it ends with a tap.
- 'circus' is another loan word in Arabic normally produced as a one-syllable word [siːɾk] (from French 'cirque'). Once again, B7's production in A combines a high front quality typical of Arabic in the first syllable, and the addition of a second syllable with a schwa typical of English. Note that the [ɪ] in ['sɪɾkəs] is short (31ms) and the /r/ is present, while the [ɜː] in ['sɜːkəs] is long (111ms) and the production is non-rhotic. B7 has therefore acquired vowel length in English in words where historic post-vocalic /r/ merged with the preceding vowel (Cruttenden, 1994: 189), but is producing a short vowel in A when followed by a tap. The same applies to the vowel in 'earth', produced with a duration of 85ms in [ɜɾθ] in A and 161ms in [ɜːθ] in E.
- The second syllables in 'singer' and 'waiter', when produced in A, have an open front quality (['sɪŋgaɾ]; ['weɪta̠ɾ]) that is typical of how the parents often produce the unstressed syllable in of the -er derivational morpheme. Note the production of (-ing)

as [ŋg] in both English and Arabic productions. It is difficult to interpret the occurrence of such realisation in B7's production. On the one hand, it might be considered as an influence from the spelling or from the way the bilinguals' parents produce the (-ing) variable in English. On the other hand, the production of [ŋg] was occasionally noted in the productions of the monolingual English children in this study and has been noted elsewhere (Stoddart et al, 1999: 76) as occasionally occurring in the speech of young speakers in Sheffield.

- A final observation on vowels concerns the first syllable in the word 'marbles', whereby the Arabic production consists of a front [a:] quality, whereas the English one is a back [ɑ:]. Note that in A, B7 produced a consonant exchange causing /r/ and /l/ to swap positions (['ma:lbər]); /r/ and /l/ substitution is typical in Arabic, especially among children still acquiring the two liquids. The realisation of 'marbles' following the exchange in A reveals further patterns that are due to Arabic influence. For instance, the /l/ is clear although it occurred in syllable final position, which is typical of /l/ production in Arabic. Moreover, while the liquid in the second syllable, regardless of the exchange, would normally be syllabic (note the production of 'marbles' in E as ['mɑ:bl̩z̥]), it is realised as [ər], therefore allowing the schwa to break the consonant-liquid cluster. A similar realisation can be found in 'purple' realised as ['pɜ:pəl] in A and vocalised in E (['pɜ:pɵ]). As mentioned for B5, schwa insertion in consonant-plus-liquid clusters in unstressed syllables is generally typical of L2 speech and was found in the parents' production.
- Another case of schwa insertion by B7 can be noted in initial consonant clusters like in 'grandma', 'fridge', and 'drumming', produced respectively as [gᵊ'rãmɑz̥], [fᵊ'rɪdʒ], and [d:ə'rəmɪŋ] in A (Figure 4.14). Attributing this to Arabic influence needs to be done in caution, because although the three consonant clusters are not permissible in Standard Arabic, the Lebanese dialect allows their production following schwa deletion (e.g. ['framto] 'I chopped it' for Standard [faˑ'ramto], [dru:ʕ] 'arms' for Standard [duˑ'ru:ʕ]).

Figure 4.14: Spectrogram of the word 'fridge' produced by B7 during the English



session (left), and then during the Arabic session (right).

- One last observation concerns the stress pattern in the words 'butterfly' and 'butter', produced with a stress on the last syllable in A ([bətəɾ̞ˈflaɪ] and [ɓʌˈtʰəɾ]), and the first one in E ([ˈbʌtᵊflaɪ] and [ˈɓʌtʰə]). Stress on the last syllable in tri-syllabic and disyllabic words occurs frequently in Arabic (e.g. [tɪffeˈħaːt] 'apples').

Table 4.12: English tokens produced by B10 in Arabic sessions compared with different productions of the same tokens in English sessions

B10	Gloss	A Produced in Arabic sessions	B Produced in English sessions
Pre-V	present (noun)	ˈpʰɾ̥ɛzɛ̃nt	ˈpʰɹ̥̆ɪzə̃nʔs
	microphone	ˈma̝ɪkɾəfõ	ˈmaɪkʰə̥fəʊ̃n
	umbrella	ʔʌmˈbrɛllə	ᵒ̊mˈbɹ·ɛlə
Post-V	beer	biɾ̞	bɪə
	circus	sɜɾˈkʰas	ˈsɜːkə̥s
	star	stɑ·ɾ	stɑːz̥
	waiter	ˈweɪta̝ɾ	ˈweɪtə
	cartoons	kʰa̝˞ˈtʰɜːn	ˈkʰɑːtʰʉːnz̥
	fireman	fɛjɛɾ̞ˈmãnː	ˈfaɪəmən
	guitar	giˈtɑ̝·ɾ̥	gɪˈtʰɑː
	deer	diːɾ	dɪə
	scarf	skɑ̝ɾ̥f	skɑ·f

As mentioned for B5 and B7, B10's production in each language session displays noticeable differences with regards to vowels, consonants, and stress patterns.

- The first vowel of the diphthong /aɪ/ in 'microphone' and 'fireman' produced in A is raised and sounds closer to the quality of an Arabic [a̝ː] found for B5 (Table 12). Moreover, the vowel in the last syllable in 'microphone' is produced with a nasalised [õ], similar to French productions of this syllable. 'microphone' is a loan word in Lebanese Arabic and is often produced the French way, [mɪkɾoˈfõ], by Lebanese speakers in general and the bilinguals' parents in particular. Note that B10's production of 'microphone' in A therefore combines a mixture of features from English, Arabic, and French.

- The vowel in 'beer' and 'deer' is monophthongal and closer to the quality of /i/ in A, while it is diphthongal and non-rhotic in E. Similarly, the first vowel in 'guitar' is closer and fronter in A than in E. As for the second vowel in 'guitar', it is short (55ms) and followed by a tap in A opposed to a longer non-rhotic [ɑː] in E (131ms). A similar pattern was observed in B7's production.
- The schwa in the second syllable of 'circus' and 'waiter' turns into a front open vowel [a] in A, similar to B7's production of comparable words and, as said before, typical of how the parents often produce the unstressed syllable of the -er derivational morpheme.
- The difference in the production of 'umbrella' in each session is interesting, as in A, the initial vowel is preceded by a glottal stop as typically found for initial Arabic vowels (though it can be found in English as well), and is realised as [ʌ] whereas in E, the quality is closer to a rounded [ʊ] that is very short but that would typical of a Yorkshire pronunciation of [ʌ]. Moreover, the /l/ is produced as a geminate in A (117ms) and a singleton in E (51ms) (Figure 4.15). The bilinguals' parents often produce words that have double 'l's in the spelling as in 'umbrella' as geminates.

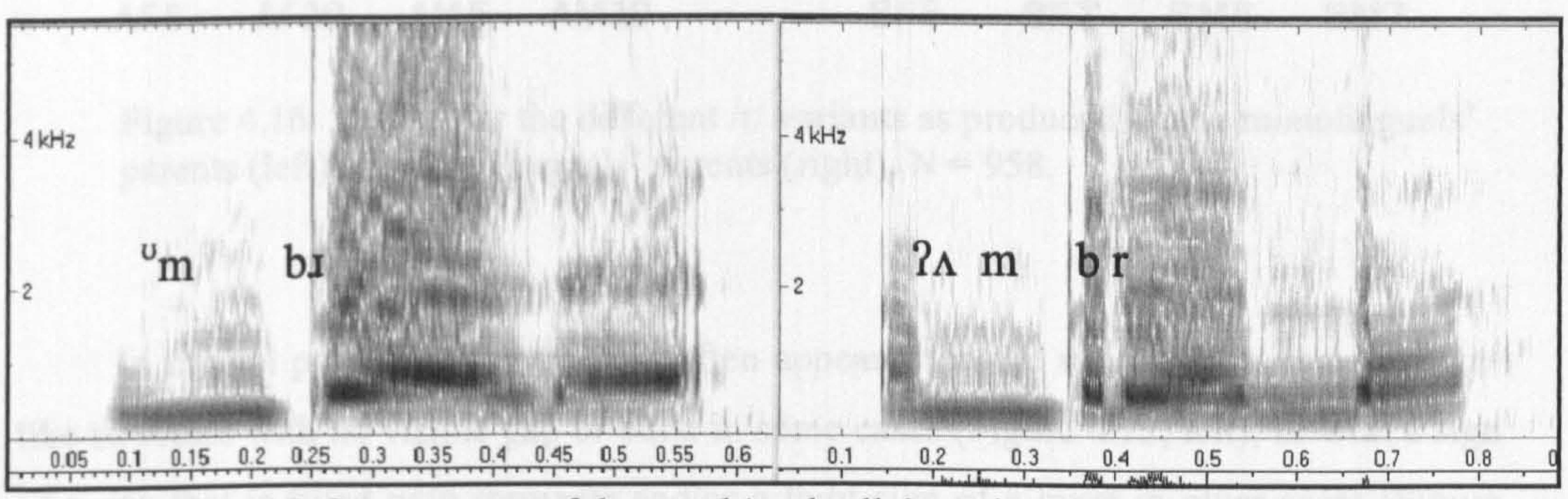


Figure 4.15: Spectrogram of the word 'umbrella' produced by B10 during the English session (left), and then during the Arabic session (right).

4.8 Arabic results

4.8.1 Adults: individual results

Figure 4.16 and Table 4.13 show the results for the /r/ variants used by each of the monolinguals' parents and the bilinguals' parents speaking Arabic. On the whole, the two groups look homogenous with respect to their /r/ production patterns, which is expected knowing that the bilinguals' parents are all native speakers of Arabic. The most frequent variant produced by all the adults is the tap. As for the weak tap and the trill, their production shows that there are idiosyncratic preferences for one type over the other, especially with regards to AM10 and BF7. Note that the occurrence of the weak tap was also influenced by syllable position, as word-final position induced more incomplete closures than other positions; however, AM10 produces weak taps frequently regardless of syllable or word position (Table 4.15). Similarly, the occurrence of the trill for BF10 did not only take place in geminates but was a frequent realization both in onsets and codas.

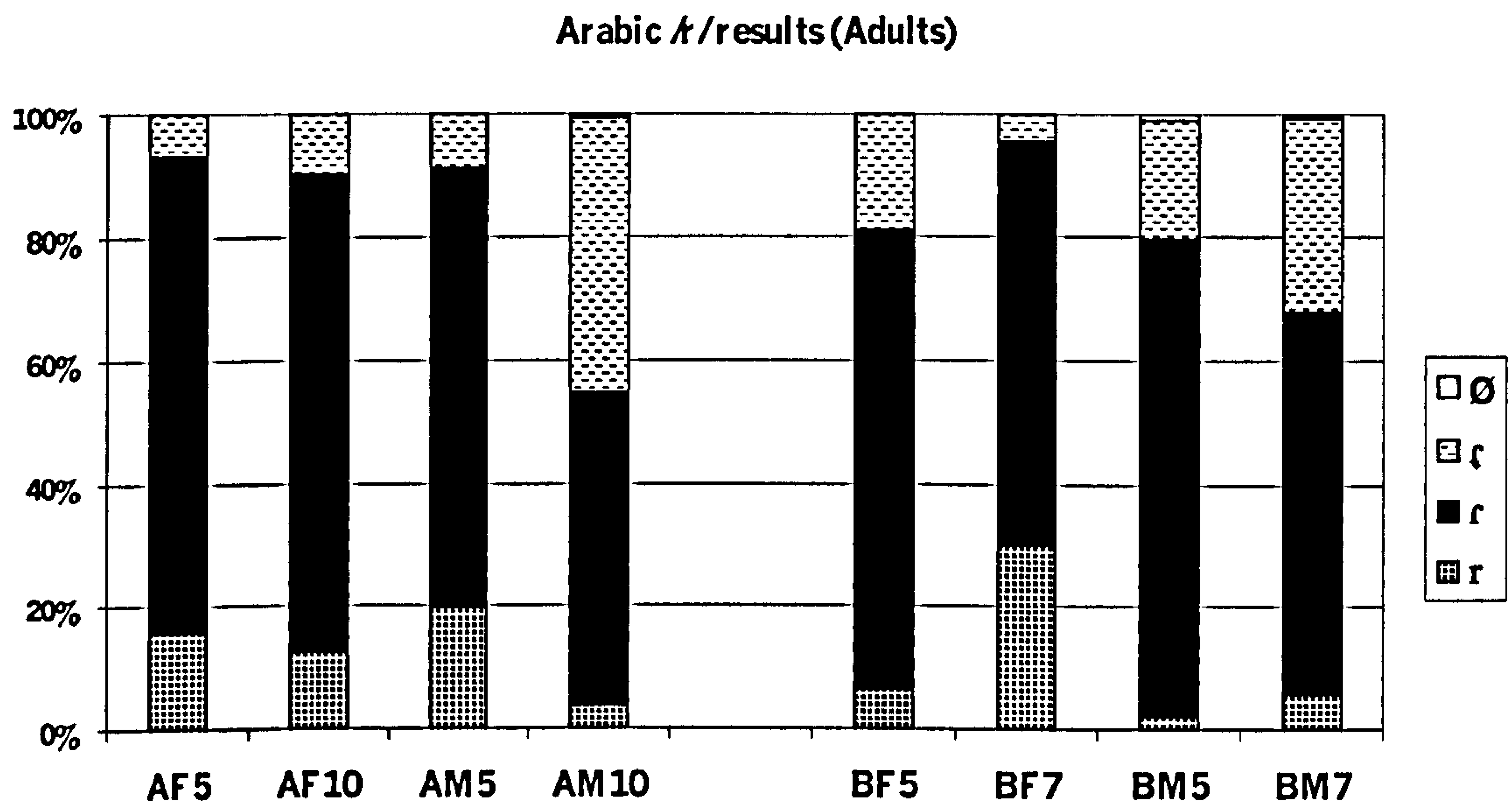


Figure 4.16: Results for the different /r/ variants as produced by the monolinguals' parents (left) and the bilinguals' parents (right). N = 958.

In medial position, the weak tap often appeared on the spectrogram as a formant-like structure with no visible gap or burst in some cases (Figure 4.17, left), or with a sign of a gap that is filled with formants and/or a faint sign of a burst in other cases (Figure 4.17, right). In final position, the weak tap appeared in the form of slight formant continuation or friction typical of a word-final tap (Figure 4.18). As mentioned before, though the auditory distinction between taps and weak taps is more or less clear, it is not always easy to distinguish between them spectrographically, and the features they show seem to operate along a continuum ranging from a strong tap articulation with a gap and a burst on the one hand, to no sign of a gap or burst on the other.

Table 4.13: Detailed figures for the /r/ variants used by the monolinguals' parents and bilinguals' parents in onset (O) and coda (C) positions.

	AF5					AF10					AM5					AM10				
	read		**story**			**read**		**story**			**read**		**story**			**read**		**story**		
	O	**C**	**O**	**C**	**N**	**O**	**C**	**O**	**C**	**N**	**O**	**C**	**O**	**C**	**N**	**O**	**C**	**O**	**C**	**N**
r	3	7	7	4	**21**	3	4	2	5	**14**	2	3	8	14	**27**	1	1	1	2	**5**
ɾ	28	16	44	15	**103**	27	13	36	11	**87**	27	19	41	11	**98**	15	11	32	8	**66**
ɾ̞	1	0	3	5	**9**	2	5	0	4	**11**	1	4	4	3	**12**	17	13	15	12	**57**
Ø	0	0	0	0	**0**	0	0	0	0	**0**	0	0	0	0	**0**	0	0	0	1	**1**
Total	**32**	**23**	**54**	**24**	**133**	**32**	**22**	**38**	**20**	**112**	**30**	**26**	**53**	**28**	**137**	**33**	**25**	**48**	**23**	**129**

	BF5					BF7					BM5					BM7				
	read		**story**			**read**		**story**			**read**		**story**			**read**		**story**		
	O	**C**	**O**	**C**	**N**	**O**	**C**	**O**	**C**	**N**	**O**	**C**	**O**	**C**	**N**	**O**	**C**	**O**	**C**	**N**
r	2	0	5	0	**7**	14	10	5	6	**35**	2	0	0	0	**2**	1	4	1	1	**7**
ɾ	28	17	25	12	**82**	21	16	26	13	**76**	25	15	26	14	**80**	28	10	27	8	**73**
ɾ̞	2	6	2	11	**21**	0	0	1	4	**5**	5	8	2	5	**20**	3	9	10	15	**37**
Ø	0	0	0	0	**0**	0	0	0	0	**0**	0	0	0	1	**1**	0	0	0	1	**1**
Total	**32**	**23**	**32**	**23**	**110**	**35**	**26**	**32**	**23**	**116**	**32**	**23**	**28**	**20**	**103**	**32**	**23**	**38**	**25**	**118**

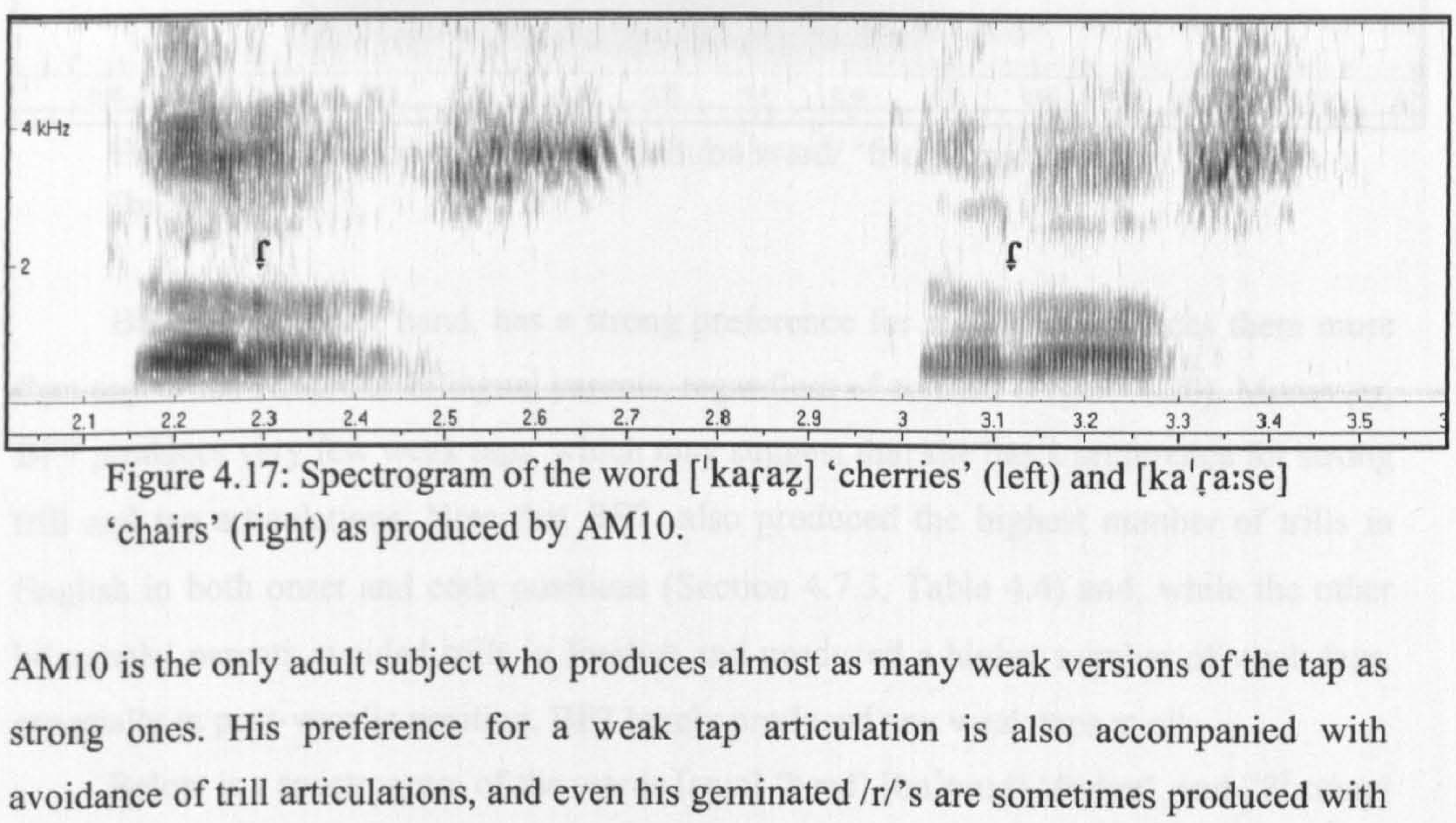


Figure 4.17: Spectrogram of the word [ˈkaɾ̞az̥] 'cherries' (left) and [kaˈɾ̞a:se] 'chairs' (right) as produced by AM10.

AM10 is the only adult subject who produces almost as many weak versions of the tap as strong ones. His preference for a weak tap articulation is also accompanied with avoidance of trill articulations, and even his geminated /r/'s are sometimes produced with one long tap that shows on the spectrogram as long filled gap (Figure 4.19).

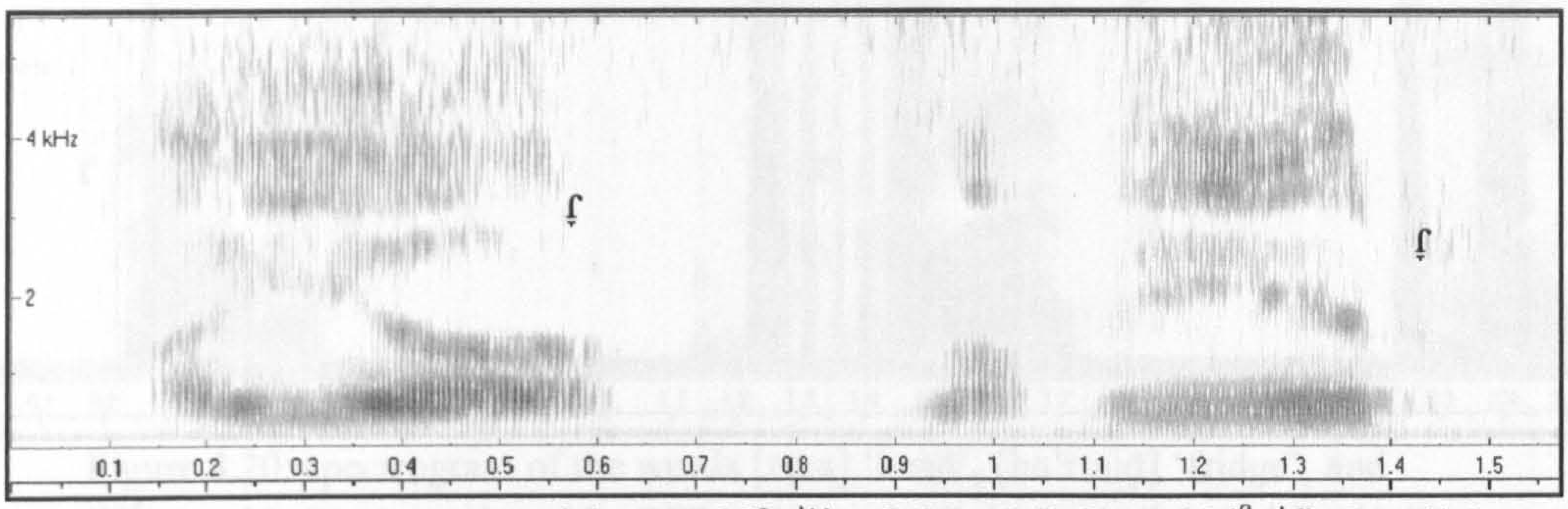


Figure 4.18: Spectrogram of the word [ʔaˈjjaːɾ̞] 'May' (left) and [dˤɑˈfiːɾ̞] 'nails' (right) as produced by AM10.

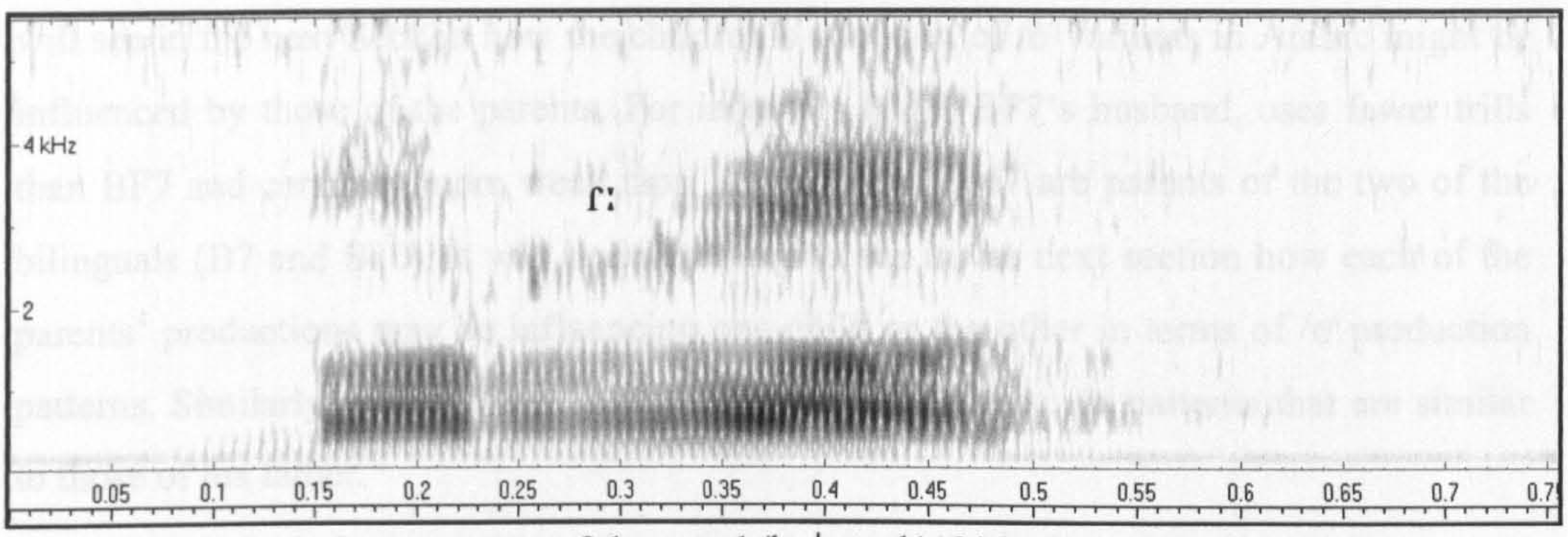


Figure 4.19: Spectrogram of the word /baˈrraːd/ 'fridge' produced by EM10 as [baˈɾːaːd].

BF7, on the other hand, has a strong preference for trills and produces them more than any of the Arabic or bilingual parents, regardless of context (Figure 4.20). Moreover, BF7 produces very few weak taps, which may suggest that she has a preference for strong trill and tap articulations. Note that BF7 also produced the highest number of trills in English in both onset and coda positions (Section 4.7.3, Table 4.4) and, while the other bilinguals' parents avoided trills in English and produced a higher number of weak taps, especially in post-vocalic position, BF7 barely produced any weak taps at all.

Below is a spectrogram of the words [rɑ̣ːs] 'head' [baˈrːaːd̥] 'fridge', and [ˈʔˤɑmar̥] 'moon', produced by BF7 and showing trill production in initial, medial, and final position. Note how different the production of [baˈrːaːd̥] 'fridge' is from that of AM10 above (Figure 4.19).

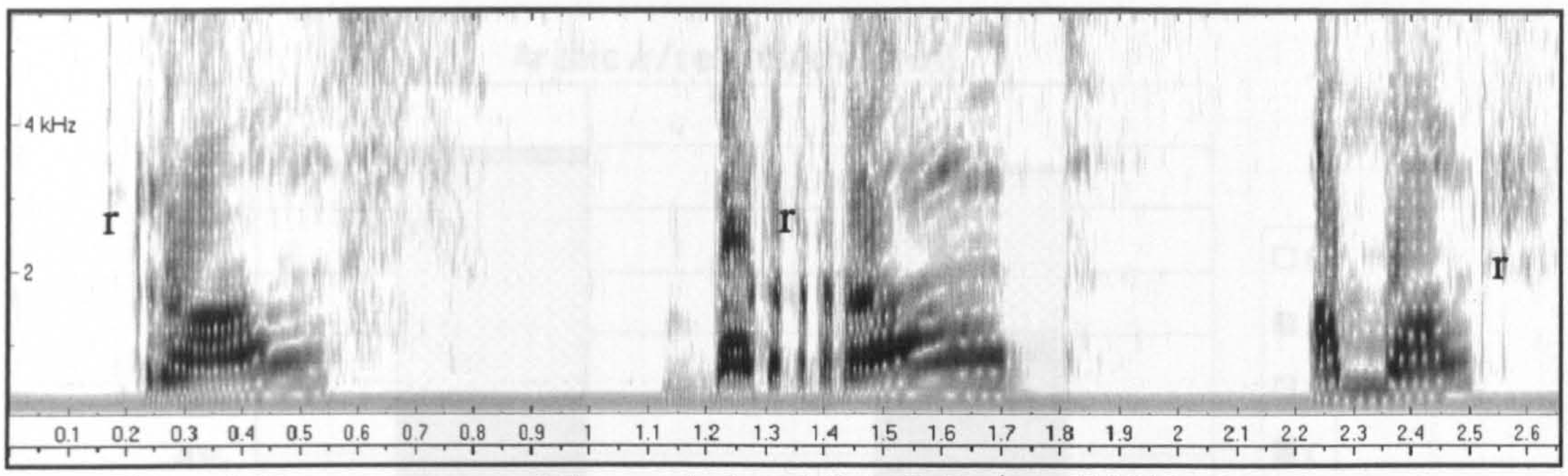


Figure 4.20: spectrogram of the words [rɑ̣ːs] 'head', [baˈrːaːd̥] 'fridge', and [ˈʔˤɑmar̥] 'moon', produced by BF7 and showing initial, medial, and final trills.

Having discussed certain idiosyncrasies about the parents' choice of /r/ pattern, we will see in the next Section how the children's adoption of /r/ variants in Arabic might be influenced by those of the parents. For instance, BM7, BF7's husband, uses fewer trills than BF7 and certainly more weak taps. As BF7 and BM7 are parents of the two of the bilinguals (B7 and B10), it will be interesting to see in the next section how each of the parents' productions may be influencing one child or the other in terms of /r/ production patterns. Similarly, EM10's son, A10 shows, as we shall see, /r/ patterns that are similar to those of his father.

4.8.2 Children: group results

Figure 4.21 shows the results for the /r/ variants used by each of the Arabic and bilingual children speaking Arabic. On the whole, the two groups seem to be using the same patterns, with the tap variants (both weak and strong) constituting the majority of the realisations, followed by trills and other realisations which will be discussed in more detail in the individual results section. The bilinguals therefore seem to have acquired the /r/ patterns for Arabic and do not show any influence from the English patterns of /r/ production. These results, coupled with the results found in Section 4.7.6 for the bilinguals' production in English, show that the bilingual subjects have acquired the expected /r/ patterns for each of their languages and are producing each set of patterns in the relevant context. Furthermore, the patterns found for the code-switched tokens reveal the importance of the social context on bilingual children's linguistic choices. This issue will be discussed in more detail in Chapter Six.

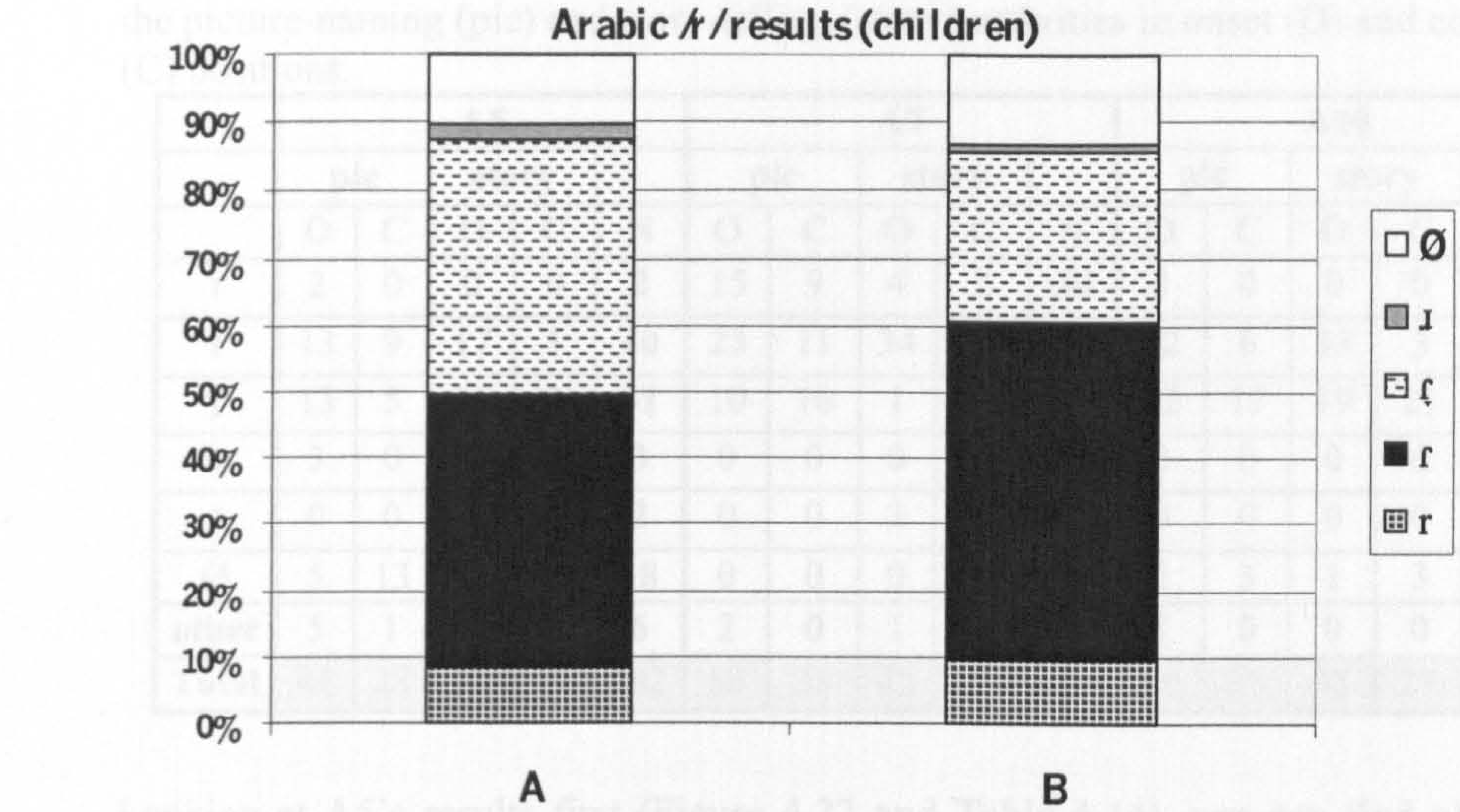


Figure 4.21: Results for the different /r/ variants produced by the monolingual and bilingual children. 'Ø' includes deletions and other realisations. N = 726.

4.8.3 Monolingual children: individual results

Starting with the results of the monolingual children (Figure 4.22 and Table 4.14), the three subjects produce a considerable number of weak taps along with strong taps, especially A10. Similarly to the results found for the parents, the children produce weak taps in all contexts and not just in final position or as a result of a fast production, strongly suggesting that [ɾ̞] should feature in the description of Lebanese /r/ alongside [ɾ] and [r].

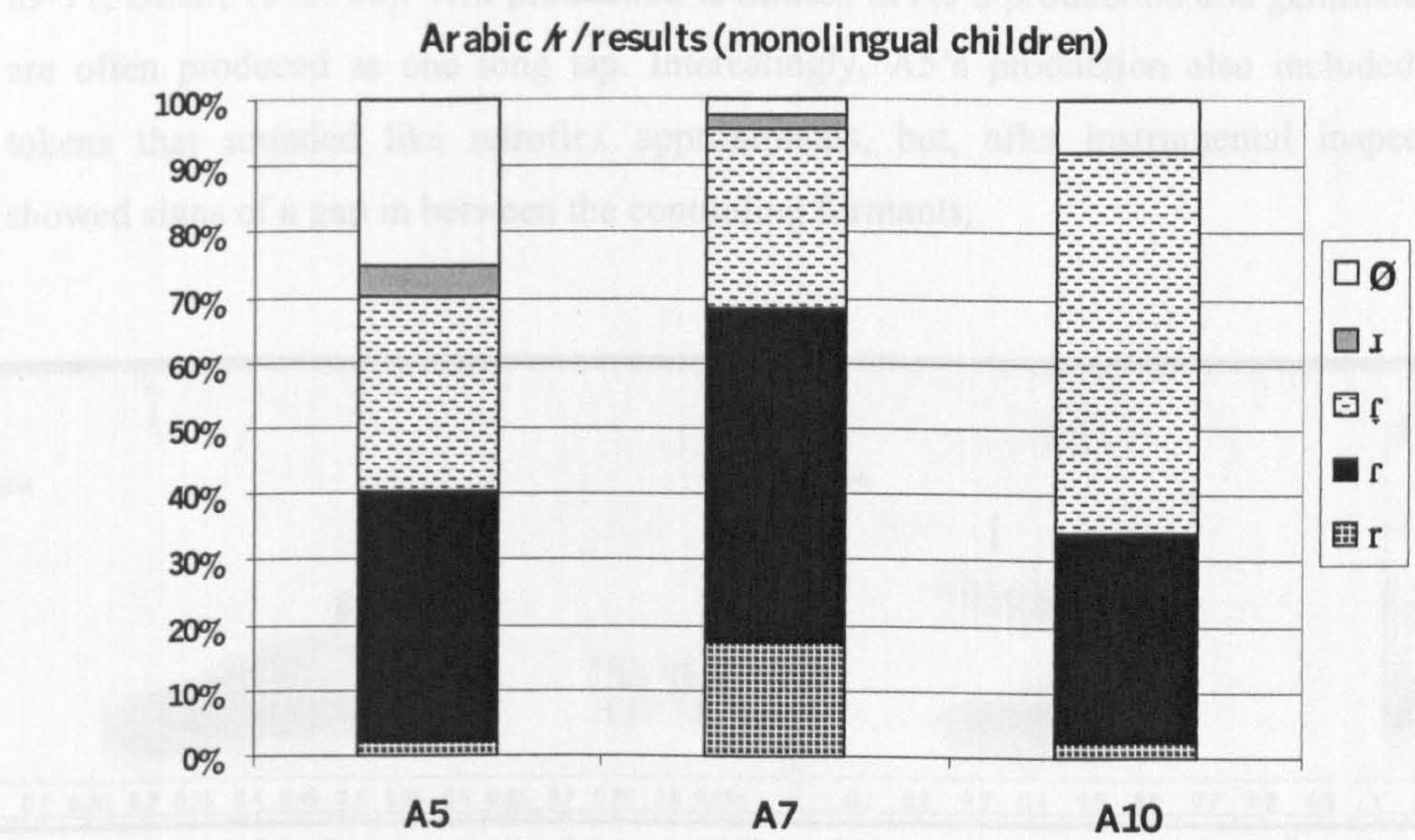


Figure 4.22: Results for the different /r/ variants produced by the monolingual. 'Ø' includes deletions and other realisations. N = 382.

Table 4.14: Detailed results for the /r/ variants used by the monolingual children during the picture-naming (pic) and story-telling (story) activities in onset (O) and coda (C) positions.

	A5					A7					A10				
	pic		story			pic		story			pic		story		
	O	**C**	**O**	**C**	**N**	**O**	**C**	**O**	**C**	**N**	**O**	**C**	**O**	**C**	**N**
r	2	0	0	0	**2**	15	9	4	2	**30**	3	0	0	0	**3**
ɾ	13	9	17	1	**40**	23	11	34	12	**80**	22	6	13	3	**44**
ɾ̥	13	5	11	2	**31**	10	16	1	13	**40**	12	17	19	21	**69**
ɹ	3	0	0	0	**3**	0	0	0	0	**0**	0	0	0	0	**0**
ɻ	0	0	2	0	**2**	0	0	2	0	**2**	0	0	0	0	**0**
Ø	5	13	0	0	**18**	0	0	0	0	**0**	2	3	1	3	**9**
other	5	1	0	0	**6**	2	0	1	1	**4**	1	0	0	0	**1**
Total	**41**	**28**	**30**	**3**	**102**	**50**	**36**	**42**	**28**	**152**	**40**	**26**	**33**	**27**	**126**

Looking at A5's results first (Figure 4.22 and Table 4.14), one can find obvious developmental features in her production in that a high proportion of her /r/'s are omitted (18%), and 6% consist of other realisations. Omissions occur mainly in word-final position, e.g. [da'f:i:] for [dˤɑ'fi:ɾ] 'nails', and ['ʔɪʒə] for ['ʔɪʒɪɾ] 'foot', but also in medial position, e.g. ['wɛ̞:ʔa̰] for ['waɾʔa] 'paper' (Figure 4.23). Other realisations consist mainly of [l] substitution of /r/, e.g. [b̥ana'd̥u:la] for [bana'du:ɾa] 'tomatoes', ['la̰ʔb̥e] for ['ɾaʔbe] 'neck', ['l:a̰ʔtˤɐ̥] for ['ɾabtˤa] 'hair band' (Figure 4.23), [kɛb'li:t] for [kɛb'ɾi:t] 'matches', and assimilation, e.g. [ma'ttaħ], for [matˤ'ɾaħ] 'place', [fɪ'ʃʃe:je] for [fɪɾ'ʃe:je] 'brush', which was also realised with a rhotic vowel [fɪ˞'ʃe:je]. /r/ lateralisation is common among children acquiring Arabic (e.g. Dyson & Amayreh, 2000: 89-91; Omar, 1973: 56). Trill production is limited in A5's production and geminate /r/'s are often produced as one long tap. Interestingly, A5's production also included five tokens that sounded like retroflex approximants, but, after instrumental inspection, showed signs of a gap in between the continuing formants.

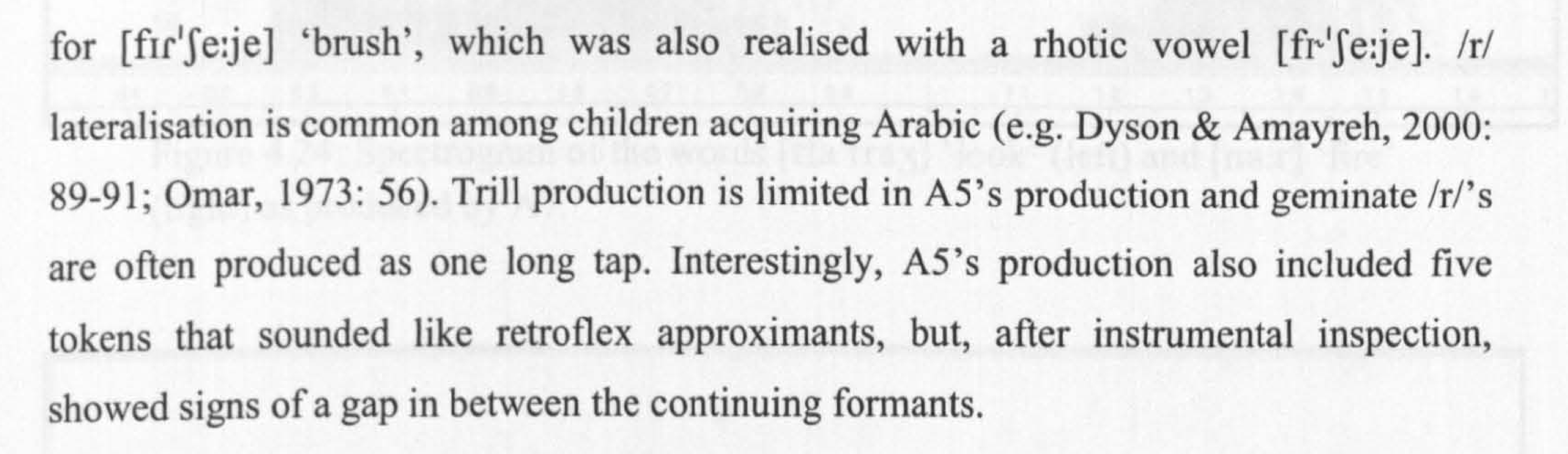


Figure 4.23: Spectrogram of the word ['waɾʔa] 'paper' (left) and ['ɾabtˤa] 'hair band' (right) realised respectively as ['wɛ̞:ʔa̰] and ['l:a̰ʔtˤɐ̥] by A5.

A7, on the other hand, did not omit any of the /r/'s he produced, but substituted /r/ in four productions including three lateralisations ([maˈbaʕlɪʃ] for [maˈbaʕɾɪʃ] 'I don't know', [lˈʔləd] for [lˈʔɑrədˤ] 'the earth', and [ˈjɪlkʊd] for [ˈjɪɾkʊd] 'he runs') and one gliding ([maˈbaʕjɪʃ] for [maˈbaʕɾɪʃ] 'I don't know'). Both types of substitutions are common in children acquiring Arabic, but there is an obvious development between A5 and A7 with regards to how often substitutions and omissions are taking place. Moreover, A7 produces more trills than A5, and in fact more than A10 as well, which may again suggest that [r] is not only a contextual variant of /r/ (occurring in geminate /r/'s), but also varies in the frequency of its production according to individual differences. 24% of A7's /r/ tokens were produced as trills regardless of syllable position (Figure 4.24). Finally, like A5, A7 produced 2 /r/ tokens that sounded like a retroflex approximant but, after instrumental inspection, showed signs of a gap or reduction in amplitude in between the continuing formants with F3 lowering (Figure 4.25).

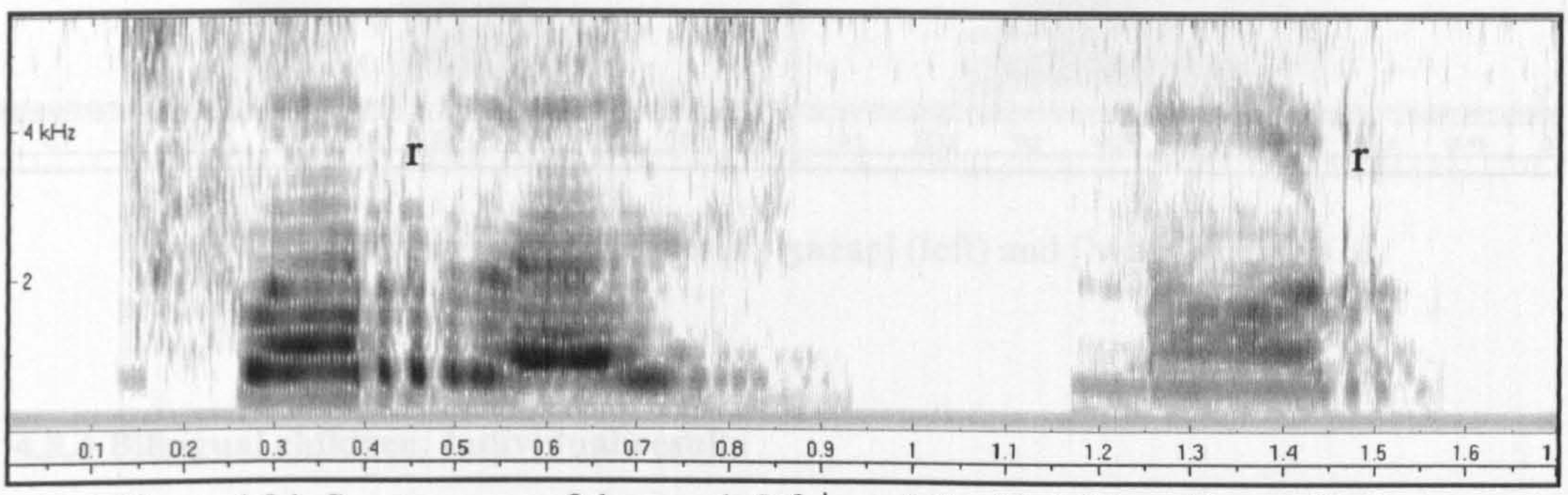


Figure 4.24: Spectrogram of the words [tfaˈrraʒ] 'look' (left) and [naːr] 'fire' (right) as produced by A7.

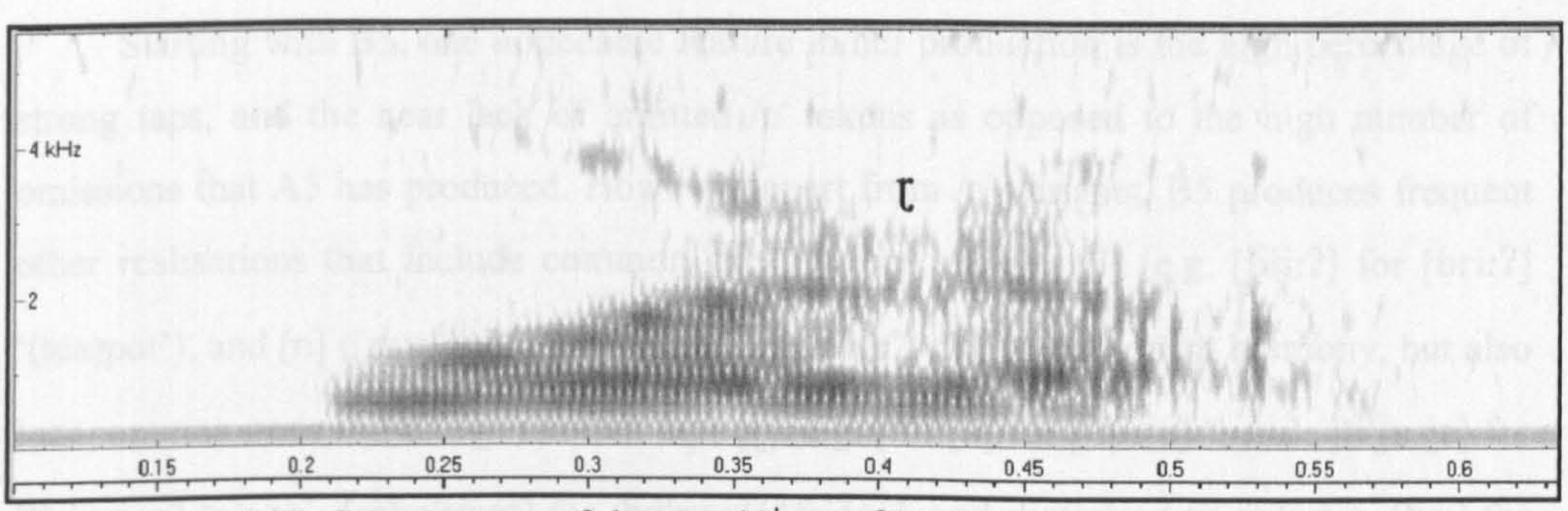


Figure 4.25: Spectrogram of the word [ˈwaɻaʔ] 'paper' produced by A7.

As for A10, the high percentage of weak taps that he produces (58%) is similar to that of his father (AM10), who also produces frequent weak realisations (44%), providing more evidence towards the suggestion that [ɾ̞] production is perhaps a feature of Lebanese

/r/ for some speakers. Note that A10 also produces a small number of trills compared with A7 (2% for A10 *versus* 17% for A7). When analysed instrumentally, most of A10's productions show as filled gaps with continued formants and, in rare cases, a faint sign of a burst (Figure 4.26). A10 also omitted nine /r/ tokens, six of which were in final position (e.g. [ˈnɪmə] for [ˈnɪmɪɾ] 'tiger'; [xja:] for [xja:ɾ] 'cucumber'), and the other three were cases of /r/ assimilation to the following sound (e.g. [tɪˈʃ:i:n] for [tɪˈʃɾi:n] 'October'; [ˈʃaʒ:a] for [ˈʃaʒɾa]). Finally, even A10 produced a token with /r/ lateralisation ([ˈmle:je] for [mɾe:je] 'mirror'), showing that such developmental features can persist even till the age of ten.

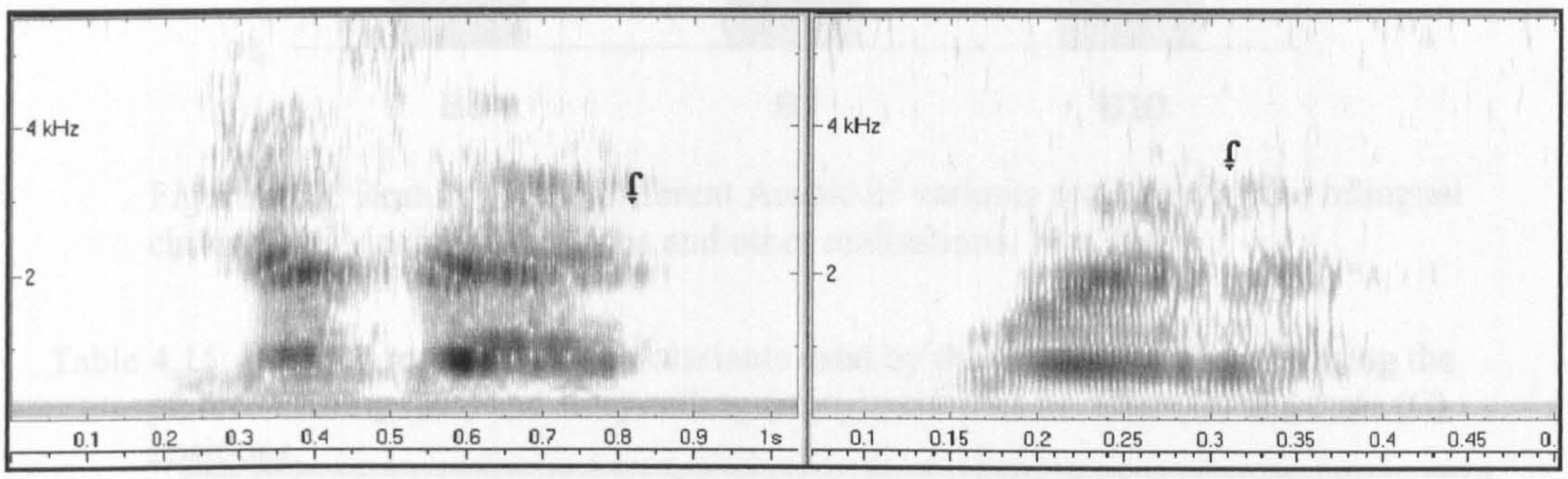


Figure 4.26: Spectrogram of the words [ˈʒazaɾ̞] (left) and [ˈwaɾ̞aʔ] (right) as produced A10.

4.8.4 Bilingual children: individual results

Figure 4.27 and Table 4.15 show the individual results for /r/ production by the bilingual children speaking Arabic. On the whole, the bilinguals' production in Arabic is similar to that of the monolinguals and there does not appear to be any influence from English.

Starting with B5, one noticeable feature in her production is the high percentage of strong taps, and the near lack of omitted /r/ tokens as opposed to the high number of omissions that A5 has produced. However, apart from /r/ variants, B5 produces frequent other realisations that include common substitutions such as [l] (e.g. [bli:ʔ] for [bɾi:ʔ] '(tea)pot'), and [n] ([mɪnˈxa:n] for [mɪnˈxa:ɾ] 'nose') due to consonant harmony, but also less common ones including retroflex taps [ɽ], e.g. [ˈbi:ɽa] for [ˈbi:ɾa] 'beer'; [tiˈjɐ:ɽɐ] for [tˤɪˈj:ɑ:ɾa] 'plane'; [mbaˈɽ:a:d] for /baˈr:a:d/ 'fridge', and rhoticised vowels e.g. [ħa˞] for [ħar] 'chillies'; [ʔɑ˞d̥ˤ] for [ʔɑɾədˤ] 'earth' (Figure 4.28). These productions sound slightly foreign accented and suggest that, although B5 is more advanced than A5 in terms of the number of omissions, she might using a wider repertoire of realizations.

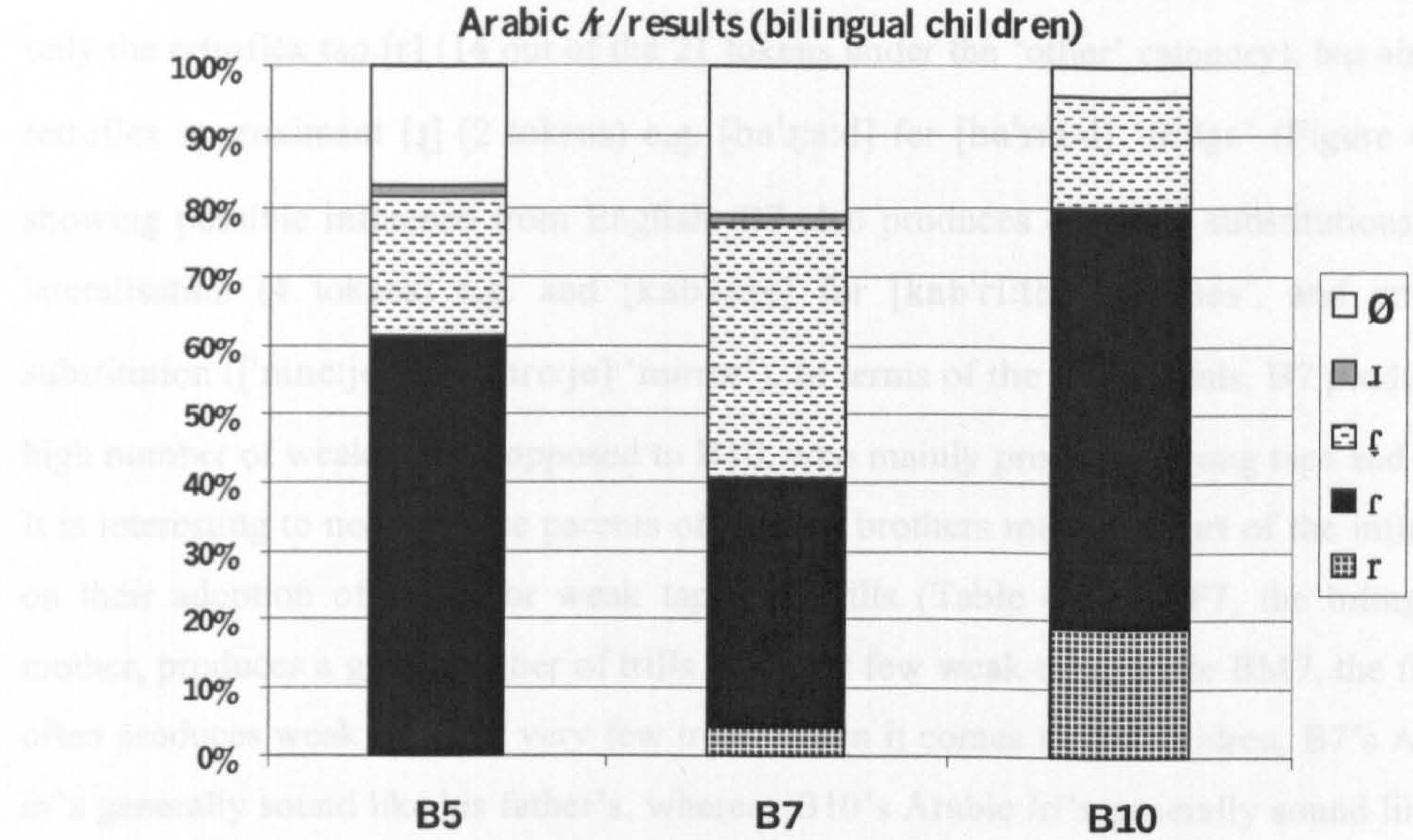


Figure 4.27: Results for the different Arabic /r/ variants produced by the bilingual children. 'Ø' includes deletions and other realisations. N = 344.

Table 4.15: Detailed results for the /r/ variants used by the bilingual children during the picture-naming (pic) and story-telling (story) activities in onset (O) and coda (C) positions.

	B5					B7					B10				
	pic		story			pic		story			pic		story		
	O	C	O	C	N	O	C	O	C	N	O	C	O	C	N
r	0	0	0	0	**0**	2	2	1	1	**6**	11	10	3	3	**27**
ɾ	22	10	6	0	**38**	23	13	14	0	**50**	26	24	28	9	**87**
ɾ̞	3	5	4	1	**13**	27	8	13	3	**51**	2	3	12	6	**23**
ɹ	0	1	0	0	**1**	1	0	0	0	**1**	0	0	0	0	**0**
ɹ̞	0	0	0	0	**0**	0	0	1	0	**1**	0	0	0	0	**0**
Ø	0	0	0	1	**1**	1	5	1	1	**8**	0	1	0	3	**4**
other	6	2	2	0	**10**	13	6	1	1	**21**	0	0	1	1	**2**
Total	**31**	**18**	**12**	**2**	**63**	**67**	**34**	**31**	**6**	**138**	**39**	**38**	**44**	**22**	**143**

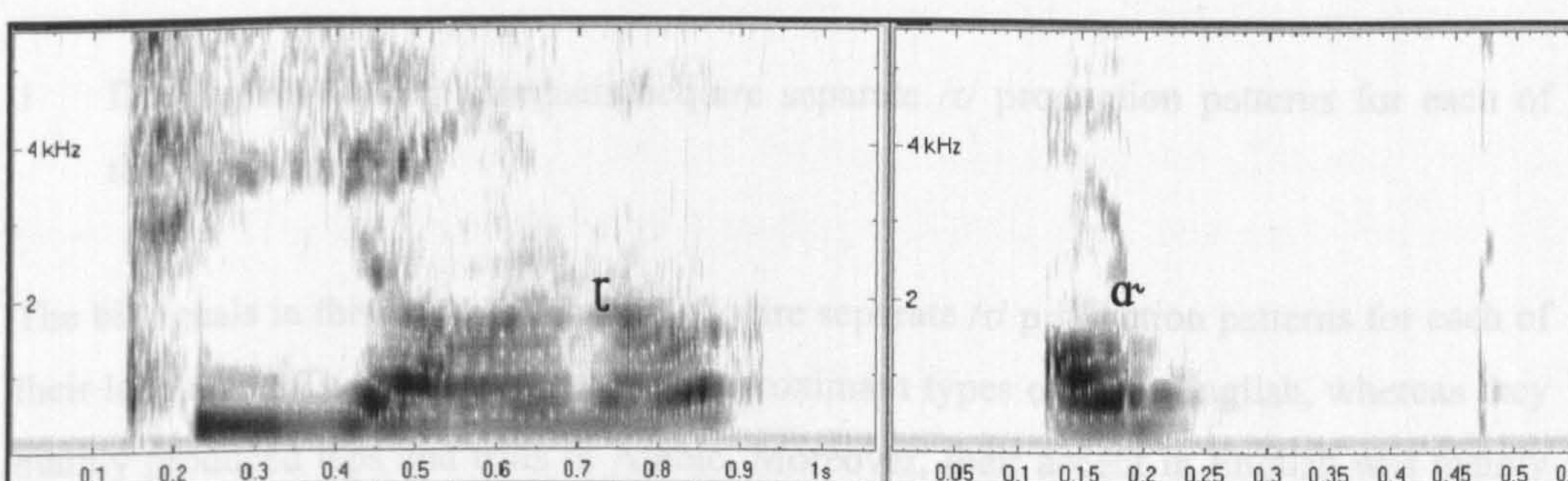


Figure 4.28: Spectrogram of the words [tˤɪˈjːɑːɾa] 'plane' (left) and [ʔɑɾədˤ] 'earth' (right) produced respectively by B5 as [tiˈjjɐɽɐ] and [ʔɑ˞d̥ˤ].

Similarly to B5, B7 produces a number of uncommon realizations including not only the retroflex tap [ɽ] (14 out of the 21 tokens under the 'other' category), but also the retroflex approximant [ɻ] (2 tokens) e.g. [baˈɻ:a:d] for [baˈr:a:d] 'fridge' (Figure 4.29), showing possible influence from English. B7 also produces common substitutions such lateralisation (4 tokens) e.g. and [kabˈli:te] for [kabˈɾi:te] 'matches', and one [n] substitution ([ˈmne:je] for [ˈmɾe:je] 'mirror'). In terms of the tap variants, B7 produces a high number of weak taps as opposed to B10, who mainly produces strong taps and trills. It is interesting to note that the parents of the two brothers might be part of the influence on their adoption of strong or weak taps and trills (Table 4.15). BF7, the bilinguals' mother, produces a great number of trills and very few weak taps, while BM7, the father, often produces weak taps and very few trills. When it comes to the children, B7's Arabic /r/'s generally sound like his father's, whereas B10's Arabic /r/'s generally sound like his mother's. B10 produced few substitutions and omissions, which suggests that /r/ patterns are more adult-like than those of B5 and B7.

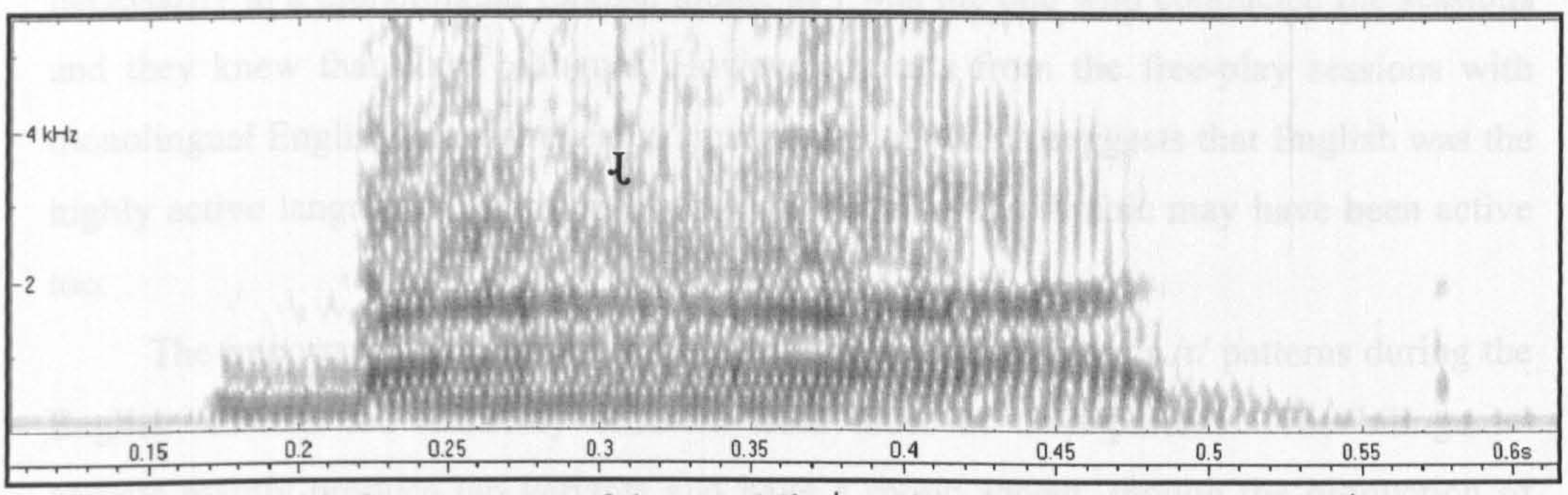


Figure 4.29: Spectrogram of the word [baˈr:a:d] 'fridge' produced as [baˈɻ:a:d] by B7.

4.9 Summary and discussion

An attempt will now be made to answer the four questions that were raised in Section 4.5.

1 Do English-Arabic bilinguals acquire separate /r/ production patterns for each of their languages?

The bilinguals in this study did indeed acquire separate /r/ production patterns for each of their languages. They mainly produced approximant types of /r/ in English, whereas they mainly produced taps and trills in Arabic. Moreover, their accent in English was mainly non-rhotic, whereas in Arabic /r/ was produced in all pre- and post-vocalic positions.

2 Are their patterns of production in each language similar to those of the monolingual controls in the study?

The patterns produced by the bilinguals were on the whole similar to those of the monolinguals. In English, both groups of children mainly produced the alveolar approximant [ɹ], while [ʋ] showed a gradual decrease with age. While in /l/ vocalisation (Chapter Three) increased with age and showed signs of being acquired as an accent feature, [ʋ] seems to be a developmental feature in the production of these children and does not show signs of being retained with age.

One minor difference between the two groups was noted in the small number of taps and post-vocalic productions that the bilinguals produced, which may be due to influence from Arabic. These do not decrease with age, and are present both in isolated word productions and running speech. Note, however, that two of the monolinguals produce sporadic taps as well. Moreover, although the recordings took place during English-only sessions and the bilinguals did not speak Arabic, the subjects were not necessarily in a monolingual English mode, as I was the one who conducted the sessions and they knew that I am bilingual. However, results from the free-play sessions with monolingual English friends revealed similar results, which suggests that English was the highly active language during those sessions, but also that Arabic may have been active too.

The important thing, however, is that the bilingual children's /r/ patterns during the English sessions are markedly different from those of their parents'. The bilinguals' parents mainly produce tap variants and have a rhotic accent, though the production of coda /r/'s is more predominant in the females than in the males. One of the males actually had a fully non-rhotic production during the story telling activity, which underlines the importance of looking at several speech styles in order to obtain a more informed idea about the relationship between the linguistic competence of L2 speakers and the task that they are involved in. Moreover, the bilinguals' parents did produce a small number of alveolar approximants and a weak variety of taps.

In Arabic, both groups of children produced mainly taps and trills, and sporadic productions of the approximant [ɹ]. Within tap production, there was a weak variant [ɾ̞] that was also found in the adults' production (see next question). This variant is normally mentioned in the literature as an approximant [ɹ] realisation of the Arabic /r/ (Shaheen, 1979), but this study has shown that there is and auditory and acoustic difference between the two types of realisations. It was therefore important to find out that this variant was not only produced by the bilinguals and therefore was not a result of influence from

English. The most frequent productions of weak taps by one of the monolinguals (A10) and one of the bilinguals (B7) actually appeared to be correlated with frequent weak tap productions by one of their parents, though more investigation is needed to confirm this observation.

Developmental features such as omissions, assimilations, and substitutions appeared in the productions of both groups of children. However, there were two minor differences between the two groups. First, developmental features in the monolingual group decreased with age whereas in the bilingual group, B7 had more omissions and other realisations of /r/ than B5. Still, B10 had the lowest number of omissions and other realisations. Second, other realisations by the monolingual included variants normally reported in the literature for children acquiring Arabic, e.g. [l], [j] and [n] realisations of /r/, assimilation to a following obstruent, etc. (Dyson & Amayreh, 2000). The bilinguals, on the other hand, produced these and other realisations not normally reported for monolingual Arabic children, including retroflex taps, retroflex approximants, and rhoticised vowels e.g. [ɚ], [ɾ], and [a˞]. These realisations show that the bilinguals have a wider repertoire of /r/ sounds than that of the monolinguals and it would be difficult to pin down the influence as coming from English, Arabic, or even other varieties that the children may be exposed to. What is important, though is that the bilinguals' /r/ patterns in Arabic are still different from the ones discussed in English on the one hand, and the English production during the Arabic sessions on the other (see question four).

3 Are the patterns for the monolingual subjects in this study similar to the ones normally described in the literature and therefore expected for each language?

In English, although there were suggestions that Yorkshire /r/ is realised as a tap (Wells, 1982), data from the IViE corpus and from monolingual English friends and their parents suggest that the alveolar approximant is the most common variant for /r/ in this small community. This in turn suggests that the taps that are reported for Yorkshire have either undergone change or are restricted to certain age groups and/or social classes. Developmental features include omission and production of a labial approximant, both of which seem to decrease as the age of the children increases. Note, however, that [ʋ] was also found to be frequent in the production of one of the monolinguals' parents who comes from London (E10) and may therefore be part of his accent. This observation, together with the patterns found for the bilinguals' parents, constitutes an example of the kind of variety in productions that bilingual children are likely to be exposed to when the parents speak English as an L2 and the families live in urban cities where the children might encounter a wide range of English accents outside the home.

Moving on to the Arabic results, data from the monolinguals' parents and the bilinguals' parents, for whom Arabic is the native language, suggest that the tap is the most common variant, followed by the trill and the weak tap, both of which proved to be interestingly related to individual preferences by the speakers. The weak tap is not mentioned in the literature, but was suggested in this study because its auditory and acoustic characteristics did not fit any of the other variants normally associated with Arabic /r/. Instrumental analysis of [ɾ] revealed a continuum of forms ranging from a filled gap with a sign of a burst like for a tap, to a formant-like structure with no sign of gap or burst and therefore closer to an approximant production, but lacking F2 and F3 lowering that are typical of English approximants.

The weak tap appeared to be frequent in some but not all of the speakers' productions. While some speakers like AM10 have a preference for weak articulations of the Arabic /r/ and produce very few trills, others like BF7 have a strong preference for trills and strong tap articulations, and produce no weak taps at all. More interestingly, some of the children's patterns suggest that they may be adopting preferences from their parents, as A10 produces a number of weak taps that is comparable to that of his father (AM10), while each of the two bilingual brothers seems to be influenced by one of the parents' productions, B10 producing strong tap and trill articulations like his mother (BF7), and B7 producing weak tap articulations like his father (BM7). More investigation of this variant is needed in order to determine whether it is correlated with gender, dialect or other social stratification. The weak tap, together with the emphatic glottal stop that was found in the production of two of the bilinguals' parents (Chapter Three), points to the need for more investigations of the phonetics and phonology of Lebanese Arabic.

4 Are there signs of influence from one language onto the other in the bilinguals' production and what are the factors that affect such influence?

Two types of influence are noted here: the first one concerns the small number of taps that were produced by the bilinguals in English, and the various types of /r/ realisations in Arabic that included common ones that were also found for the monolingual Arabic controls but also less common ones. This first type of influence was minimal and did not show a great deal of interaction between the two languages.

The second type of influence concerns the bilinguals' English productions during the Arabic sessions. As opposed to the /r/'s produced during the English sessions, the majority variant used for the English tokens from the Arabic sessions is the tap, along with a considerable number of post-vocalic productions and only a small proportion of the alveolar approximant. Figure 4.30 combines the results for /r/ patterns by the bilinguals from the controlled and free English sessions, the Arabic sessions, and the English produced during the Arabic sessions. The results from the three contexts suggest that different language modes were operating in the bilinguals during each context, with obvious overlap. This issue will be discussed in more detail in Chapter Six.

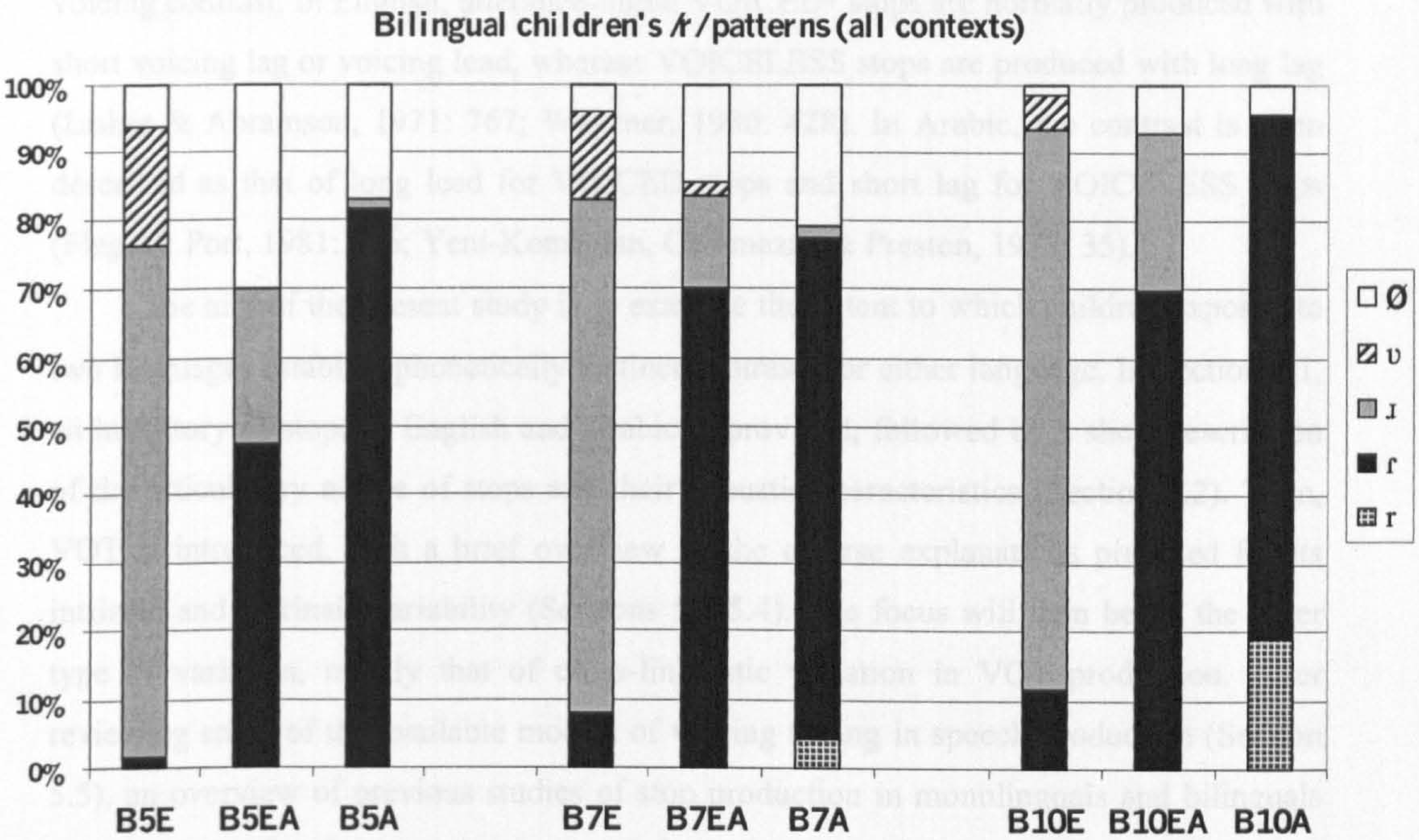


Figure 4.30: Summary of the /r/ patterns found for the bilinguals in the three different language contexts (E = English; EA = English in Arabic context; A =Arabic).

CHAPTER FIVE

Voice Onset Time

5.0 Introduction

This chapter presents an investigation into the Voice Onset Time (VOT) patterns developed by the bilingual subjects for their production of English and Arabic initial stops. English and Arabic vary considerably in their phonetic realisation of the stop voicing contrast. In English, utterance-initial VOICED[9] stops are normally produced with short voicing lag or voicing lead, whereas VOICELESS stops are produced with long lag (Lisker & Abramson, 1971: 767; Weismer, 1980: 428). In Arabic, the contrast is often described as that of long lead for VOICED stops and short lag for VOICELESS stops (Flege & Port, 1981: 126; Yeni-Komshian, Caramazza & Preston, 1977: 35).

The aim of the present study is to examine the extent to which children exposed to two languages establish phonetically distinct contrasts for either language. In Section 5.1, an inventory of stops in English and Arabic is provided, followed by a short description of the articulatory nature of stops and their acoustic characteristics (Section 5.2). Then, VOT is introduced, with a brief overview of the diverse explanations provided for its intrinsic and extrinsic variability (Sections 5.3-5.4). The focus will then be on the latter type of variation, mainly that of cross-linguistic variation in VOT production. After reviewing some of the available models of voicing timing in speech production (Section 5.5), an overview of previous studies of stop production in monolinguals and bilinguals sets the stage for the investigation carried out in my study, that of the production of Arabic and English monolingual and bilingual subjects (Sections 5.6-5.7). Details of the experiment including the aims, subjects, and procedure are presented in Sections 5.8 and 5.9, while results of the VOT patterns emerging in each language will be presented and interpreted in Sections 5.10-5.12, taking into consideration the age and linguistic background of each of the speakers. The findings offer a contribution to existing research on bilingual phonological acquisition as well as an up-to-date profile of VOT patterns in English and Arabic.

5.1 English and Arabic stops

Table 5.1 shows the stop inventories of English and Lebanese Arabic taken from Davenport & Hannah (1998) and Nasr (1966) respectively:

[9] From this point onwards, following the convention used in Docherty (1992), capital letters will be used for the terms 'VOICED' and 'VOICELESS' to refer to their phonological status, and small letters 'voiced' and voiceless' to refer to their physical status.

Table 5.1: Stops in English and Arabic

	Bilabial		Dental		Alveolar		Post-alveolar		Velar		Glottal
English	p	b			t	d			k	g	
Arabic	(p)	b	t	d			$t^{ʕ}$	$d^{ʕ}$	k	(g)	ʔ

In Arabic, no native /p/ exists. However, proper names and loan words, principally from French (*piscine, pyjama*), are frequent in the Lebanese dialect and are usually produced with [p] by the majority of people, especially educated ones. [b] and [b̥] can also be heard as realisations for [p], especially among the uneducated. Similarly, no native /g/ exists in Lebanese Arabic, but people usually produce it accurately in loan words (*garage, gateau*). One reason might be that /g/ is familiar to the Lebanese due to its use by nearby Arabic dialects such as Palestinian and Egyptian (e.g. Al-Shareef, 2002). However, [k] is sometimes heard as a realisation for /g/ in loan words. For the purposes of this chapter, only /p t k/ and /b d g/ will be examined in each language. /$t^{ʕ}$/ and /$d^{ʕ}$/ will not be included in the analysis because they occur infrequently in the children's speech and because the main aim is to compare the bilinguals' VOT production in each language.

5.2 Nature of stops and their acoustic characteristics

Stop segments can be produced on a range of phonation varying from complete voicelessness to strong voicing[10], depending on their phonological identity but also on the surrounding context (e.g. Borden and Harris, 1984: 120; Kent & Read, 1992: 106-110; Laver, 1994: 340) and the language in question. Stops normally consist of three physical events: (i) a closure phase (onset phase), in which an active articulator moves to contact a passive articulator; this can be detected visually on a spectrogram by the presence of particular formant transitions in vocalic sounds preceding the closure; (ii) a hold phase, in which the closure is maintained and air pressure builds up behind it; this is detected on a spectrogram by the presence of an acoustic gap and silence in the case of voiceless stops, while voiced ones exhibit energy at a low frequency only; (iii) a release phase (offset phase), in which the constriction is released, air begins to flow at high speed, and there is an immediate burst of energy occupying a wide range of frequencies. This shows on the spectrogram as a vertical transient and is usually called the release burst (Kent & Read, 1992: 106-110; Laver, 1994: 340).

In words with prevocalic initial stops (CV(C) pattern), a fourth event may occur if the burst is followed by some turbulent noise energy, as there is a brief period during

---

10 Other phonation types like breathy, whisper, creak, etc. are not discussed here.

which the articulators are still close enough to cause friction, and, in the case of VOICELESS stops, the glottis is still partially open. This phenomenon is known as aspiration, and it usually appears on the spectrogram as aperiodic energy, usually in higher frequencies. The noise may show similarities with that of the glottal fricative [h], thus the superscript [ʰ] in the IPA representation of aspiration. The degree of aspiration varies depending on the degree of glottal opening during the closure, i.e. the greater the opening, the longer the amount of aspiration (e.g. Ladefoged, 1982: 132), the rate/volume of airflow, as well as the place of articulation of the stop and the quality of the following vowel.

Since in some languages like English aspiration plays a distinctive phonological role, the difference in phonological terms between a set of VOICELESS stops and a set of VOICED ones is not usually just one of phonation during the consonant closure. Instead, both sets of sounds may be realised as unphonated, and the phonological difference may be signalled by the presence or absence of aspiration. Thus, a distinction needs to be drawn between the phonological terms VOICED/VOICELESS, and the terms aspirated/unaspirated referring to the state of the glottis during a given articulation and to the presence or absence of voicelessness before and after the release of an articulation.

5.3 VOT: a definition

Voice Onset Time (VOT) is a term that was coined by Lisker & Abramson (1964) in their classic cross-linguistic study of phonation in initial stops in eleven languages. The authors defined VOT as 'the time interval between the burst that marks the release of the stop closure and the onset of quasi-periodicity that reflects laryngeal vibration' (Lisker & Abramson, 1964: 422). They conducted their experiment in order to test how well VOT serves to separate the phonological stop categories in a number of languages. The measure of VOT was found to be highly effective in the languages examined, although these differed in the number of phonological categories and in the phonetic features assigned to them.

Physiological and aerodynamic factors such as place of articulation of the stop, position in the syllable, and speech rate all play a role in the timing of voicing (see discussion in Section 5.4.1). VOT, however, is not an inevitable consequence of these factors. The timing of the events must be learned during the acquisition of the grammar of any language (see Section 5.6). Evidence lies in the way languages select different targets along the VOT continuum (see. Section 5.4.2). Due to a combination of internal and external factors affecting VOT, Cho and Ladefoged's (1999) more recent VOT definition will be adopted in this study, since the authors give an active role to VOT production in terms of the voluntary initiation of gestures by speakers for the realisation of a particular timing for vocal fold vibration. VOT is defined as 'the time between the *initiation* of the articulatory gesture responsible for the release of a closure and the *initiation* (my emphasis) of the laryngeal gesture responsible for vocal fold vibration'[11] (Cho and Ladefoged, 1999: 225).

There is a continuum of possible duration of the time difference between the release and the onset of voicing, which constitutes a physical scale along which the realisation of stops can be located. The most important finding that emerged from Lisker and Abramson's study is that there are three VOT categories that delimit the glottal and supraglottal relationships for the stop systems of many languages (Figure 5.1). VOT can thus be assigned to three types of values: (i) a negative VOT, or voicing lead, occurs when phonation begins before the release burst; a typical fully voiced stop has a VOT of approximately -60ms; (ii) a VOT of zero value occurs when phonation starts simultaneously with the release burst. Stops with a VOT of 0 to 30ms are either VOICELESS unaspirated stops or VOICED stops with no initial voicing during the closure period; (iii) a positive VOT, or voicing lag, occurs when phonation is delayed after the release burst. A typical voiceless aspirated stop has a VOT of approximately 60ms.

There are, however, finer divisions within each of these three phonetic categories that are equally important in distinguishing between stop categories that are often considered 'similar' in two given languages with regards to the phonetic category they occupy along the VOT continuum. Evidence from Cho & Ladefoged (1999) and from this study will be presented later in the chapter.

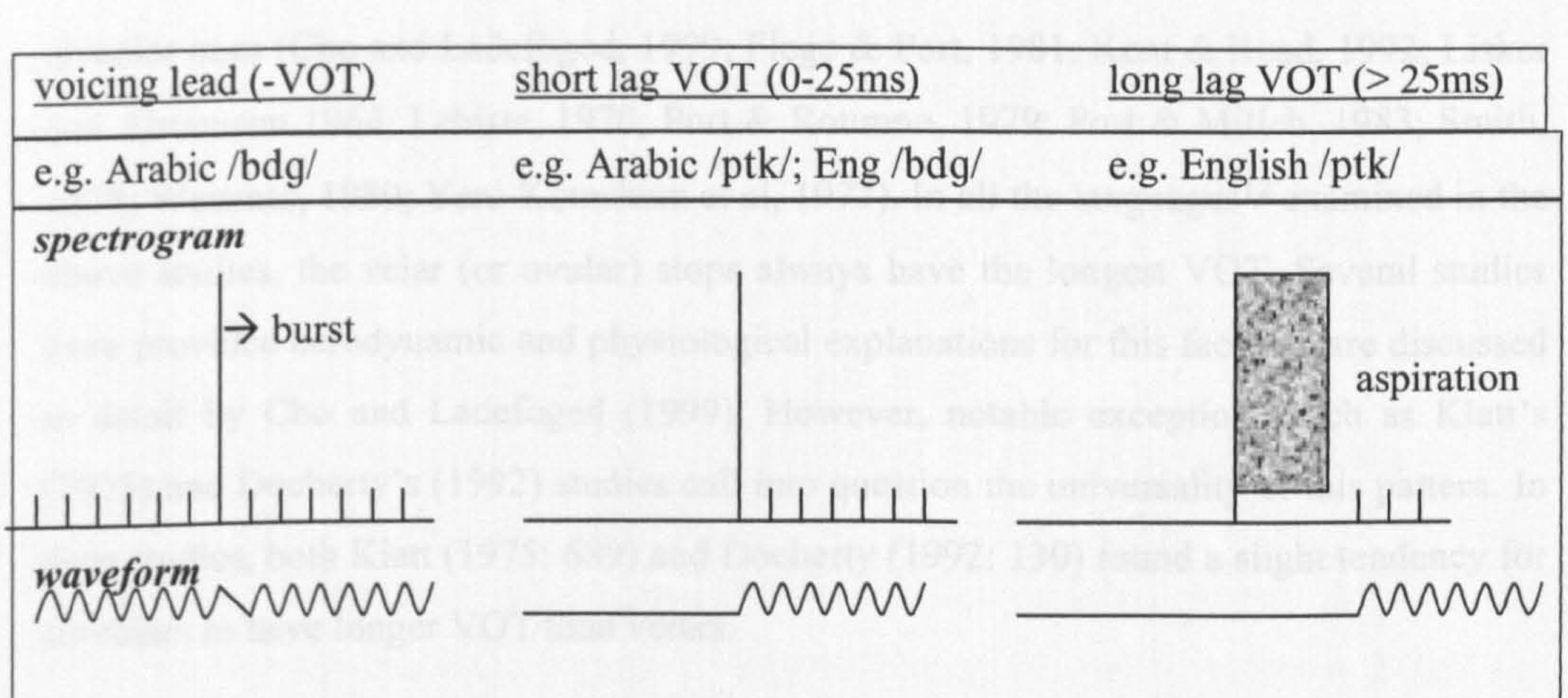


Figure 5.1: Schematic representation showing the relation between the release burst and vocal fold vibration in three phonetic categories: voicing lead, short lag, and long lag.

[11] Though in terms of acoustic measurements, Cho & Ladefoged (1999) use the same method to measure VOT as Lisker & Abramson (1964).

5.4 Universal and language-specific variations in VOT

5.4.1 Universals in VOT

There are fairly universal parameters that manifest themselves in languages with respect to the timing between the glottal and supraglottal activities required for stop production. This uniformity often relates to inherent properties of sounds (e.g. place or manner of articulation), properties of the vocal organs (physical, mechanical, and inertial), the influence of the linguistic context in which the sounds find themselves (e.g. position in the syllable, number of syllables in the word, sentence position, quality of the following vowel, interarticulator coordination), and more global temporal and prosodic factors such as stress and speech rate (e.g. Docherty, 1992: 20; Lehiste, 1970: 18).

One such universal is that VOT for a voiceless stop is longer before close vowels than before open ones. This is due to the fact that high tongue body position for close vowels offers greater resistance to the outflow of air from the vocal tract, thus delaying to a greater extent the onset of airflow of sufficient volume for vibration of the vocal cords to occur (Laver, 1994: 353; Catford, 1977: 197). The effect of close vowels on VOT has been acknowledged in many studies, including Flege & Port (1981), Jesry (1996), Radwan (1996), Smith (1978), and Yeni-Komshian et al (1977).

However, it is often difficult to differentiate between actual intrinsic constraints and language specific influences on the manipulation of the timing of stops, since they are subject to continuous variation among a number of parameters, and within which there are no hard boundaries. For instance, another pattern that is often reported to be universal is the variation of VOT as a function of the different places of articulation of the stop: VOT in VOICELESS stops tends to be longer for velar stops than for bilabial and alveolar ones (Cho and Ladefoged, 1999; Flege & Port, 1981; Kent & Read, 1992; Lisker and Abramson 1964; Lehiste, 1970; Port & Rotunno, 1979; Port & Mitleb, 1983; Smith, 1978; Weismer, 1980; Yeni-Komshian et al, 1977). In all the languages[12] examined in the above studies, the velar (or uvular) stops always have the longest VOT. Several studies have provided aerodynamic and physiological explanations for this fact and are discussed in detail by Cho and Ladefoged (1999). However, notable exceptions such as Klatt's (1975) and Docherty's (1992) studies call into question the universality of this pattern. In their studies, both Klatt (1975: 689) and Docherty (1992: 130) found a slight tendency for alveolars to have longer VOT than velars.

---

[12] Dahalo is the only exception in Cho & Ladefoged's study, where alveolar stops have longer VOT than velar ones. However, the authors offer an articulatory explanation for this phenomenon, suggesting that it is due to an unusually slow articulation used for the alveolar stops in this language compared to the velar ones.

Moreover, results from languages that have uvular stops such as Arabic also show disagreement with the general tendency of VOT to increase as the place of articulation moves further back, partly due to the complex history of Arabic /q/ and the possibility that it has undergone a process of change from voiced to voiceless (Ibn Khaldun, 1958). For instance, Jesry (1996) and Radwan (1996) found longer VOT values for /k/ than for /q/ in their studies on Syrian Arabic, while Yeni-Komshian, Caramazza, & Preston (1977) found longer VOT values for /q/ in their study on Lebanese Arabic. Since there are numerous exceptions to the apparently universal rule, such differences in patterns might be a reflection of systematic language- and dialect-specific variability. This type of variability will be discussed in the following section.

5.4.2 Language-specific variation in VOT

Since Lisker and Abramson's study, a considerable number of investigations that have taken place seem to confirm the generality of the use of VOT as an important acoustic parameter in the production and perception of stops in many languages (e.g. Caramazza & Yeni-Komshian, 1974; Deuchar & Clarks, 1995; Flege & Port, 1981; Hazan & Boulakia, 1993; Keating, Linker, & Huffman, 1983; Klatt, 1975; Simon & Fourcin, 1978; Smith, 1978; Stevens & Klatt, 1974; Yeni-Komshian et al, 1977; Watson, 1995).

In most studies, VOT is acknowledged as being effective in separating homorganic stop categories and as being under the speaker's control. In other studies, however, VOT is considered to be a redundant feature that is 'predictable' from phonetic context, or that is the 'by-product' of dichotomous laryngeal behaviour resulting from the interaction between glottal and supraglottal articulations on the one hand, and structural and contextual factors on the other (Flege & Hammond, 1982; Weismer, 1979). This traditional view of VOT variation is based on the idea that languages select zones of comfortable phonetic performance within the intrinsic constraints on the duration of their segments to allow for relative ease of articulation and security of distinction of contrastive sounds (cf. discussion by Laver, 1994: 433).

However, two characteristics are important when discussing language-specific variation. The first is that variation may be systematic across different speakers' performance, rather than being unpredictable or in free variation. Recent investigations show that the speaker can control the coordination of the timing between glottal and supraglottal articulations, as well as the degree of glottal opening, which is a crucial factor in the timing of voicing. The second is that variation is not necessarily distinctive, meaning that it is not crucially involved in ensuring the systemic distinctions of the language or accent concerned (Docherty, 1992: 59). Speakers learn sets of patterns appropriate to their accent during language acquisition, including patterns which are sociolinguistically relevant, regardless of whether those patterns are contrastive or not. Three types of studies have been useful in offering evidence for the existence of language-specific variation with respect to VOT. The first type is that of large-scale studies that are able to control for non-systematic variation that results from using different types of measurement and elicitation techniques. Examples include three comprehensive studies, one by Lisker and Abramson (1964), who investigated VOT in 11 languages, the second by Keating, Linker, & Huffman (1983) who investigated VOT in 51 languages, and the third by Cho and Ladefoged (1999), who investigated VOT in 18 languages representing 12 language families. All three investigations show clear patterns of language-specific temporal variation that could not have resulted purely as the by-product of motor execution stages.

The second type of study is one that examines the performance of non-native speakers of a given language. There is ample evidence in the literature to show that second language learners often substitute the fine phonetic details of the target language with those of their native language, even when they succeed in realising the voicing contrast in the target language (e.g. Flege, 1980; Port & Mitleb, 1983; Riney & Takagi, 1999).

The third type of study is one that examines the performance of bilingual speakers, especially where the two languages being acquired differ in their VOT patterns. A successful acquisition of the VOT patterns that are specific to each language is prime evidence of the voluntary initiation of gestures by bilingual speakers for the realisation of a particular timing for vocal fold vibration depending on the language being used. The three types of studies will be reviewed in more detail in Sections 5.5-5.6.

5.5 Monolingual VOT studies

5.5.1 VOT studies on English

There are many reports in the literature on VOT patterns for adults in English, including Chen (1970), Docherty (1992), Keating, Mikos, & Ganong (1981), Klatt (1975), Lisker & Abramson (1967), Port & Rotunno (1979), Scobbie (2002), Smith (1978) and Weismer (1979), but Docherty (1992) and Scobbie (2002) are two of the few reports on VOT in British English. Following are some of the results obtained for VOT in isolated word-initial position.

Following the pioneering study in which Lisker & Abramson (1964) showed that VOT measures can serve as an effective basis for distinguishing between homorganic stop categories in many languages, the authors suggested the need for a closer look at the individual languages to give a more detailed picture of the relations between stop voicing and other features of each language. In 1967, they conducted a more detailed study of voicing in American English stops, and measured VOT of word-initial /p t k/ and /b d g/ in isolated words and running speech. For the isolated condition, four subjects read around 500 words that were mainly monosyllables. While /p t k/ had VOT distributions that fell mainly in the long lag range, /b d g/ showed values that fell into two discontinuous ranges, with modes about –100ms and near 0ms. Two means were therefore extracted for the VOICED set. Table 5.2 shows the mean VOT values obtained for the four speakers whose stops were analysed:

Table 5.2: Mean VOT values for stops in isolated word-initial position in Lisker & Abramson's study (1967: 6).

Stop	p	t	k	b	d	g
VOT(ms)	58	70	80	1 -101	5 -102	21 -88

Though /b d g/ values occupy discontinuous ranges, a single speaker is responsible for 95% of the stops produced with voicing lead, while another speaker is responsible for the remaining ones, which suggests that each speaker nearly always produced a single kind of /b d g/. Further experiments with those two speakers showed that they are more likely to voice their stops in isolated words rather than in sentence condition, and when asked to read minimal pairs rather than a random list. The authors suggested that the production of voicing lead by those two subjects, apart from being part of their idiolect or dialect, was being used mainly to enhance the contrast between the two stop categories (Lisker & Abramson, 1964: 24).

On the whole, the values shown in Table 5.2 suggest that there are two distinct categories for VOICED and VOICELESS stops in isolated position with no overlap. Other relevant factors that were later discussed as having an effect on VOT include stress, number of syllables, and position in the sentence. Though the values that are later presented for the sentence condition are significantly shorter due to temporal compression in running speech (Lisker & Abramson, 1964: 10), the authors maintain that it is only a case of reduction of the gap separating the distinct categories but that there is no serious overlap. They also suggest that there must be another sub-set of acoustic cues that reflect the opening or closing of the larynx and that serve to distinguish /p t k/ from /b d g/ in each context in which the contrast is applied in the language.

Klatt (1975) measured VOT of word-initial plosives preceding vowels and in consonant clusters in English, and tested the effect of place of articulation of the stop and following vowel or sonorant. Three subjects read monosyllabic target words all beginning with different word-initial clusters and single plosives embedded in the frame sentence 'Say __ instead'. As expected, VOT varied according to place of articulation and vowel context, and was considerably reduced when the plosives were preceded by /s/, but increased considerably when the plosives were followed by sonorant consonants due to a lower first formant in the following sonorant (e.g. /tr/-/kl/). Table 5.3 shows the results obtained by Klatt (1975) for VOICED and VOICELESS plosives followed by a vowel.

Table 5.3: Mean VOT values for stops in isolated word-initial position in Klatt's study (1975: 689).

Stop	p	t	k	b	d	g
VOT (ms)	47	65	59	11	17	27

Prevoiced tokens were ignored in the presentation of the results because prevoicing was not considered important for phonemic distinction in English (Klatt, 1975: 688). The study differs from other studies in that VOT was measured from the beginning of the release burst till the onset of vertical striations in the second and higher formants rather than in F1. VOT was also divided into two sections, frication and aspiration, so that the acoustic characteristics of each section could be studied in more detail. Some overlap was found in the VOT values for VOICED and VOICELESS plosives, which suggested that a perceptual decision about the voicing feature could not be made on the basis of VOT alone. Klatt (1975: 695) suggested the use of five other acoustic cues for the perception of voicing in English, including low frequency energy in following vowels, burst loudness, fundamental frequency, segmental duration, and prevoicing.

While the studies reviewed above use American varieties of English, the most comprehensive one on British English is Docherty's (1992) study of the timing of voicing in English obstruents. This study is one of the few attempts at examining various aspects of the fine detail of voicing timing in VOICED and VOICELESS obstruents in Southern British English (SBE) and at evaluating the patterns observed within a general model of speech production. VOT and voiced intervals of stops and fricatives were measured in the speech of five adult male speakers in a variety of contexts (different positions in the word, adjacent vocalic/consonantal sound, word/sentence condition). Table 5.4 shows the mean VOT values obtained for stops in word-initial position (negative tokens were not included).

Table 5.4: Mean VOT values for stops in isolated word-initial position in Docherty's (1992) study

Stop	p	t	k	b	d	g
VOT(ms)	45.74	66.45	66.09	25.00	32.84	39.96

Despite the fact that the results show a significant effect of the underlying phonological category in relation to VOT, there was an overlap between the distribution for VOICED and VOICELESS stops in all subjects, which indicates that in physical terms there is no simple binary pattern which correlates with the phonological terms VOICED and VOICELESS (Docherty, 1992: 116). Moreover, like Lisker & Abramson (1967), Docherty (1992) found two distinct VOT patterns for VOICED stops in post-pausal syllable-initial position: (i) prevoicing resulting in negative VOT values ranging between –19 and –143ms; and (ii) voicing lag with values ranging between 0 and 52ms. Though negative tokens were produced by only two of the five subjects in the study and in some environments more than others (labial and dental), they constitute a problem for models based on two distinct phonetic categories for VOICED and VOICELESS stops.

Scobbie's (2002) study underlines the importance of considering dialectal differences with respect to VOT production. Although the subjects in the study were monolingual English adults, their bidialectal background triggered a VOT acquisition process that is similar to that found in some bilingual situations (reviewed in Section 5.6). The subjects were 12 young adults (aged 16-30) who were born and raised in Shetland, and whose parental accents ranged between Shetlandic/Shetlandic (four subjects), Shetlandic/English (four subjects), and Shetlandic/Scottish (four subjects). The Shetlandic pattern for VOT is that of prevoicing for VOICED stops and short lag for VOICELESS stops, which is similar to the pattern found in other languages like Spanish, French, and Arabic. English and Scottish, on the other hand, follow the prevoicing/short lag pattern for VOICED stops, and long lag for VOICELESS stops. Table 5.5 shows the mean VOT values obtained for stops in word-initial position.

Table 5.5: Mean VOT values for stops in isolated word-initial position in Scobbie's (2002) study

Stop	p	t	k	b	d	g
VOT(ms)	56.00	66.00	75.00	-29.00	-25.00	-6.00

Although the pooled results for VOICELESS stops are similar to those found for the English studies reviewed above, the ranges were very wide and the values stretched over both the short and long lag regions. There was a considerable amount of overlap between the values for VOICED and VOICELESS stops, the former ranging from voicing lead to short lag, and the latter from short lag to long lag. Individual results showed that some subjects were using three categories for their stops (prevoicing, short lag, and long lag), while others were only using two (prevoicing/ short lag or short lag/long lag). Only one subject, whose parents were both Shetlandic, produced short lag values for both VOICED and VOICELESS stops, though the values for VOICED stops seemed to occupy lower ranges than those for VOICELESS stops. Scobbie (2002) noted that, although VOICED and VOICELESS stops might occupy slightly different ranges within the short lag category, there is a great amount of overlap and other cues should be considered for the voicing contrast. More importantly, the author draws attention to the fact that the kind of dialectal and cross-speaker variation that were found in the study are actually the norm for most monolingual and bilingual speakers. This kind of variation should therefore be taken into consideration by researchers working on child language acquisition.

5.5.2 Child studies

Reports on child VOT patterns include Foulkes et al (1999), Gilbert (1977), Kewley-Port & Preston (1974), Macken & Barton (1980), Simon (1976), Smith (1978b), Snow (1997), Stoel-Gammon & Buder (1999), and Zlatin & Koenigsknecht (1976). The usual development seems to be for all stops to be initially produced in the short lag range during the early acquisition stages. By 24 months, VOT distinctions usually start to emerge, and the production is extended to the long lag and long lead ranges. Children are also known to produce VOT with longer duration and more variability than adults do. Adult-like consistency is usually achieved around 10-12 years of age, after reductions in the duration of speech sounds and in variability gradually have taken place. Still, there are important individual differences in the developmental patterns of children, and gradual decrease in the duration of sounds is not always the norm (cf. Smith & Kenny, 1999). Below are some of the results obtained from the child studies mentioned above.

Some of the early data available on VOT production by children before the age of two can be found in Macken and Barton's study (1979), in which the authors examined the acquisition of voicing in word-initial stop consonants by children aged about 1;6 to 2;4. The children first produced predominantly short lag stops, but had generally acquired an adult-like contrast by the age of 2;0. The authors also conducted a similar study on Spanish monolinguals (Macken & Barton, 1980) where the pattern found was different from that of the monolingual English children. While the Spanish children first produced predominantly short lag stops as in the English data, the Spanish lead/lag voicing contrast did not develop as early as the English contrast. Few tokens of voicing lead stops were found in the data for the two-year-olds, and even the four-year-olds had not fully developed the lead/lag contrast. Most of their tokens fell in the short lag range, and the evidence of voicing contrast that they used was often based on short lag for VOICELESS *versus* continuants for VOICED. The idea that voicing lead develops late is supported by other studies on Spanish and French (Allen, 1985; Konefal & Fokes, 1981). This is due to the fact that the stop closure results in rapid rise in intra-oral air pressure and a progressive balance of trans-glottal pressure leading to the difficulty in maintaining voicing (Docherty, 1992: 62), especially for children, as they have small mouth volumes.

Another study on the early acquisition of speech timing in English is that by Snow (1997), who followed the speech development of ten English-speaking girls aged between 1;6 and 2;0. The aim of the study was to compare the acquisition of segmental features such as VOT distinction with that of suprasegmental ones such as final-syllable vowel lengthening (FSVL) in order to find out which one will be acquired first and what developmental and/or language-specific factors affect such acquisition. Data collection started when the productive vocabulary of the children was at least 30 words, and lasted for about nine months when the children's vocabulary reached about 70 words. The subjects were taped in semi-structured play activities centred around toys, and their productions were grouped into one of three categories depending on the relevant period of the child's development: single-word utterances, multi-word utterances, and beginning of syntax. Two criteria were used to determine whether the children had acquired the voicing contrast: when they consistently used longer VOT for VOICELESS targets than for VOICED ones, and when the VOICELESS targets had a VOT value of more than 60ms.

Results showed that the children had acquired the first criterion from the earliest stage (single-word utterances), and the second criterion by the second and third stages. There were of course individual differences in the rate of acquisition, and all the children but one had acquired the contrast by the end of the study. The acquisition of FSVL, however, started developing at a later stage (at the beginning of syntax), and achieved significance only when compared to earlier stages. Eight out of the ten children acquired VOT before FSVL, and only one child acquired FSVL first. While the rate of VOT acquisition was largely related to the acquisition of expressive vocabulary regardless of word-combinations, FSVL acquisition seemed related to combinatorial speech.

This order of acquisition, however, does not apply to all languages. A similar study conducted with French children aged between 1;9 and 2;8 found that they acquired FSVL earlier than VOT, mainly due to articulatory difficulty associated with the gestures required for voicing lead in French (Allen, 1985). The children had shown signs of contrast acquisition, but produced their VOICED stops with short lag rather than the long lead found in adult production. They also tried to use other devices to signal the phonemic voicing distinction, one of which was to precede VOICED targets with a nasal or vowel segment that permitted continuous voicing.

Stoel-Gammon & Buder (1999) examined aspects of speech timing in the production of 20 two-year-old children acquiring American English. The main measures included VOT in word-initial stops, extrinsic and intrinsic vowel duration of tense/lax high vowels in CVC syllables, and voicing of final obstruents. On the whole, 50% of the children's productions fell within an appropriate range for maintaining a voicing contrast for word-initial stops, with 10 out of the 20 subjects showing signs of stable acquisition of the voicing contrast (with at least 75% accuracy for both VOICED and VOICELESS targets). Children who performed well on this feature exhibited reliable use of extrinsic vowel lengthening as well, which suggested that they were more phonetically advanced than the other children in the study. The authors expected that VOICED stops would be produced correctly (due to the fact that they fall in the short lag range in English), and that difficulties would occur with the VOICELESS targets. However, since nearly 90% of the initial stops that were analysed preceded high front vowels, the VOT measures obtained from this study tended to be longer than those reported for children producing words with vowels more evenly distributed in terms of vowel height. As a result of that, low accuracy rates were found for VOICED stops, which tended to be produced with a VOT exceeding the expected range.

One of the few studies available on VOT acquisition of British English is by Foulkes et al (1999), who conducted an investigation of the speech of 40 children aged 2 to 4 from Newcastle upon Tyne. The aim of the study was to understand how several phonetic patterns that are particular to the children's community are acquired by those children, including the production of (t), a variable known to be complex in adult speech (Docherty & Foulkes, 1999). With respect to the production of word-initial /t/, preliminary analysis of the productions of 10 children showed that they had all mastered the production of the long lag variant expected for English in this position. Although the subjects were still variable in terms of the wide VOT ranges they produced, all but three out of the 215 tokens analysed fell in the long lag VOT category (longer than 25ms). Mean VOT values were between 63 and 134ms, which conforms with general observations that segmental durations are longer in children's speech.

At this point it is worth noting that the lack of consistency in producing adult-like VOT values at an early age does not necessarily mean that the child cannot perceive the phonological contrast in the adult language. Judging the children's production abilities using cues that are salient in the adult production patterns might lead to missing other important cues that the children might be using to achieve voicing contrast. For instance, in their investigation of the voicing in the production of monolingual children with phonological disorders, Scobbie, Gibbon, Hardcastle & Fletcher (2000) found that one of their subjects was reported as failing to produce the voicing contrast in words beginning with /t/, /d/ and /st/ by neutralising the VOT values for the three stops in comparable contexts (all three had VOT values in the short lag region). However, this child was actually producing a 'covert contrast', i.e. one that is not recorded in transcription, by successfully manipulating other acoustic cues in the production of these stops, mainly the steepness of spectral tilt immediately following voice onset. The child had steeper spectral slope after voiceless stops than after voiced ones (Scobbie et al, 2000: 205). Since steeper spectral tilt can be found in the speech of some adults and is achieved by prolonging breathiness at the onset of vowel phonation after aspirated stops, the authors concluded that their subject had acquired some of the relevant motor skills needed to convey the contrast in VOICED and VOICELESS initial stops and was therefore showed awareness of the relevant phonological contrast, but still needed to master other language-specific phonetic skills.

5.5.3 VOT studies on Arabic

There are a handful of studies on VOT in a variety of Arabic dialects (Al Ani, 1970; Al Ghamdi, 1990; Flege & Port, 1981; Jesry, 1996; Port & Mitleb, 1983; Radwan, 1996; Yeni-Komshian, Caramazza, & Preston, 1977). Although these studies vary in their methodologies and results, they offer useful evidence on the type of phonation commonly found in VOICED and VOICELESS Arabic stops. They all agree on the fact that Arabic uses prevoicing for VOICED stops and short lag for VOICELESS ones (although some results show slight degrees of aspiration). Some of these studies will now be reviewed briefly.

Al Ani (1970) measured the duration of aspiration in his own production of VOICELESS stops in Iraqi Arabic. The tokens were uttered in words in isolation. The measurements obtained for /t/ ranged between 30-40ms, while those for /k/ ranged between 60-80ms.

Flege & Port (1981) investigated the phonetic implementation of the stop voicing contrast in word-initial stops in the Saudi Arabian dialect. Six adult males read Arabic word lists inserted in the carrier sentence /ʔagɾa __ wamʃɪlelbeːt/ 'I read __ and then I go home'. Unlike most studies done on Arabic, the sentences were produced in the colloquial Arabic of Saudi Arabia rather than Standard or Classical Arabic. The test words consisted of /CV:C/ minimal pairs, and measurements of VOT and vowel duration were made, along with the closure interval of initial and final stops. Table 5.6 shows the mean VOT measurements obtained from this study.

Table 5.6: Mean VOT values (in ms) for word-initial stops in Flege and Port's study (1981) on Saudi Arabian Arabic

Stop	t	k	b	d	g
VOT (ms)	37	52	-85	-82	-75

/t/ and /k/ were found to be slightly aspirated in this dialect, with the values for /t/ ranging between 20-65ms and those for /k/ between 30-85ms. The closure intervals of VOICED stops were produced with glottal pulsing in 100 % of the cases for /b/ and /d/, and 92% of the cases for /g/. More interestingly, a few /t/ and /k/ tokens were also produced with glottal pulsing during the closure period, which indicates a small degree of overlap between VOICED and VOICELESS tokens. In addition to a VOT difference, the durations of the stop closure intervals of VOICELESS stops were significantly longer than those of their VOICED counterparts.

Similar results were later obtained from another study of VOT in Saudi Arabic by Al Ghamdi (1990), although the subjects were speakers of the Ghamdi dialect. The author measured VOT in stops in all positions in the word. Four subjects read target words in isolation. Table 5.7 gives the VOT results for stops in word-initial position from his study.

Table 5.7: Mean VOT values (in ms) in Al Ghamdi's study (1990) on the Ghamdi dialect

Stop	t	k	b	d	g
VOT (ms)	32.32	42.12	-72.04	-71.09	-68.7

In 1996, two studies comparing the voicing contrast in English and Syrian Arabic obstruents were conducted at Essex University. In the first one, Jesry (1996) undertook a cognitive approach to voicing by seeing VOT as a coarticulatory effect that is controlled at a 'cognitive phonetic' level. In his approach, cognitive rules provide the phonetic system with the ability to produce a set of controlled articulations that do not overlap, and to control variability to improve discrimination within the available space (Jesry, 1996: 82). The author taped three Syrian adult speakers reading word lists in Modern Standard Arabic (MSA) in the sentence frame /qa:la ... ala:n/ *he said ... now*. The target words contained Arabic stops and fricatives in word-initial position followed by one of the Arabic vowels /i, i:, a, a:, u, u:/. Closure duration and VOT were measured for all tokens. Table 5.8 shows the mean VOT values for some of the stops in all vowel contexts. Although the mean VOT value for /k/ is longer than for /t/, there was no significant difference between the two. However, no category overlap was found between VOICED and VOICELESS stops, since they were separated by an interval of 80ms.

Table 5.8: mean VOT values (in ms) in Jesry's study (1996) on Syrian Arabic

Stop	t	k	b	d
VOT (ms)	27.82	32.19	-68.72	-66.8

In the second study, Radwan (1996) also taped three male Syrian subjects reading words in isolation in the carrier sentence /qa:la ... mɪra:ran/ *he said ... repeatedly* in MSA. The target words had stops in word-initial and word-medial positions followed by long or short Arabic vowels. VOT for VOICELESS stops was measured from the release burst to the onset of vertical striation in F2 (rather than F1). Similarly to Jesry's (1996) results for Syrian subjects, Radwan found that VOICED stops are characterised by a predominance of glottal pulsing during the entire closure interval. Table 5.9 shows some of the results for stops in isolated words for all three speakers:

Table 5.9: mean VOT values (in ms) in Radwan's study (1996) on Syrian Arabic

Stop	t	k	b	d
VOT (ms)	33.57	38.81	-71.03	-78.23

The only study involving adult Lebanese speakers is one conducted by Yeni-Komshian et al (1977), who investigated VOT production and perception in word-initial stop consonants by eight Lebanese subjects aged between 16 and 34, although the elicitation technique involved asking the informants to read material in MSA. The authors wanted to investigate the generality of the effect of place of articulation of the stop and the vocalic context on VOT production. They also wanted to find out whether VOT was sufficient to distinguish between homorganic sets of stops in Arabic. Production tests consisted of reading words in isolation and a text containing 21 target words in MSA. Perception tests were also carried out and consisted of asking the subjects to imitate a set of synthetic CV syllables in which VOT values were manipulated.

Results from production tests agreed with other studies reviewed in this section, in that they showed VOICED stops to be characterised by a predominance of a voicing lead. The VOICELESS stops, as in some of the other Arabic studies, fell in the short lag range. Table 5.10 shows the mean results for all 8 speakers:

Table 5.10: Mean VOT values (in ms) for some of the stops in Yeni-Komshian et al's study (1977) on Lebanese Arabic

Stop	t	k	b	d
VOT (ms)	25.00	28.33	-65.00	-56.66

The VOT ranges for stops with a voicing lead were broader than the ones for the short lag stops. More importantly, there was an overlap for all subjects between the VOT ranges of the homorganic pairs /t d/ and /tˤ dˤ/, and some of the /b/ values were in the short lag range. The authors concluded that there must be other cues besides VOT that may serve to distinguish the pairs of sounds.

Results from perception tests were interesting because they were not only in the form of responses to forced choices about the identity of stops as in many other studies, but included imitation responses that reflected the subjects' perception of 'acceptable' stops in Lebanese Arabic. Whenever the synthetic stimuli had VOT values that exceeded the short lag or long lead ranges normally expected for Arabic, the subjects failed to identify the consonants in the CV syllable as stops, and reproduced them as fricatives (/h/ in the case of long lag and /z/ or /ð/ in the case of long lead). All the other imitations that were identified as stops showed a change in perception from VOICED to VOICELESS stops at locations on the VOT continuum that were consistent with the production results, but there was considerable intra- and inter-subject variation.

Although the above studies were made with speakers from a variety of Arabic dialects, all but Flege & Port (1981) use data elicitation techniques that involve reading words or sentences in MSA rather than the vernacular of the informants. Moreover, all of the studies use careful laboratory speech rather than natural speech data. The present investigation, on the other hand, uses more natural data elicitation techniques that require the subjects to produce words in their dialects and whereby the recordings took place in the subjects' homes (see Chapter Two).

Still, several important outcomes from the studies reviewed above are relevant to the present study. First, they all show that VOT is important for distinguishing homorganic sets of stops in word-initial position in Arabic, but the studies that found some overlap in VOT measures between the set of VOICED and VOICELESS stops maintain that there must be other important cues as well, especially for the distinction between the minimal pairs /t d/ and /tˁ dˁ/. Second, in all the studies, VOICED stops are produced with predominant prevoicing during the closure period, while the VOICELESS stops range between short lag and slight aspiration. Third, in all the studies, VOT for the velar VOICELESS stops is longer than for the alveolar ones, although the difference is often insignificant. However, the universal rule about VOT being longer as the place of articulation moves further back in the mouth does not apply to all the studies when the measurements for /q/ and /tˁ dˁ/ are considered, as there are complications related to emphasis and historical changes in /q/, but these issues will not be discussed further in this study. As for the VOICED stops, there is no clear pattern for VOT in terms of place of articulation of the stop. Finally, all the studies acknowledge the prevalence of intra- and inter-speaker variations, the latter type occurring mainly due to individual differences such as speech rate.

5.5.4 Child studies

There are hardly any studies on the acquisition of VOT by Arabic children, let alone Lebanese ones. The only study I am aware of is by Preston et al (1967), who attempted a cross-cultural comparison of apical stop production in one Lebanese and one American infant who were both 12 months old. The authors found that both infants produced their stops in the short lag region with VOT values ranging between 0 and 30ms, and concluded that short VOT intervals may be the easiest for infants to accomplish as opposed to voicing lead and long lag which require careful timing between supraglottal and glottal articulators. Similar cross-linguistic evidence has been provided by Enstrom (1982) from Swiss-German data, and Macken & Barton (1980) and Kewley-Port & Preston (1974) from English data.

5.5.5 Summary of English and Arabic VOT patterns

Table 5.11 shows a summary of the mean VOT values found by some of the English and Arabic studies reviewed above.

Table 5.11: Summary of mean VOT values in ms for stops in isolated word-initial position found for English and Arabic.

English	p	t	k		b	d	g
Lisker & Abramson (1967)	58.00	70.00	80.00	+	1.00	5.00	21.00
				-	-101.00	-102.00	-88.00
Klatt (1975)	47.00	65.00	59.00	+	11.00	17.00	27.00
Docherty (1992)	45.74	66.45	66.09	+	25.00	32.84	39.96
				-	ranged between -143 and -19		
Scobbie (2002)	56.00	66.00	75.00		-29.00	-25.00	-6.00

Arabic		t	k	b	d	g
Flege & Port (1981)		37.00	52.00	-85.00	-82.00	-75.00
Al Ghamdi (1990)		32.32	42.12	-72.04	-71.09	-68.70
Jesry (1996)		27.82	32.19	-68.72	-66.80	
Radwan (1996)		33.57	38.81	-71.03	-78.23	
Yeni-Komshian et al		25.00	28.33	-65.00	-56.66	

Although the English and Arabic studies reviewed in this section vary in their focus and results, some common generalisations can still be made about VOT patterns in English and Arabic. These are summarised in Table 5.12.

Table 5.12: VOT patterns in word-initial position in English and Arabic

English	Arabic
- Presence or absence of vocal fold vibration in the closure duration of stops is not contrastive.	- Presence or absence of vocal fold vibration in the closure duration of stops is contrastive.
- Initial VOICELESS stops are characterised by a delay between 50 and 80ms in voicing relative to the release of the stop.	- Initial VOICELESS stops are characterised by a delay between 25 and 60ms in voicing relative to the release of the stop.
- Initial VOICED stops are either unaspirated (VOT is between 0 and 25ms) or voiced	- Initial VOICED stops have a predominance of voicing lead (VOT is between –60 and –90ms).

While both English and Arabic fall into the two-category group of languages in terms of the number of stop categories they contain (Lisker & Abramson, 1964: 388), the two languages vary considerably in their patterns of phonetic implementation of the stop voicing contrast. Voicing contrast in English and Arabic is associated with a number of acoustic cues and is more complicated than the usual description of the binary opposition of VOICED/VOICELESS. For instance, in English there need not be any vocal fold vibration during the production of either of the pairs /p b/, /t d/, /k g/. However, /p t k/ tend to be aspirated, while /b d g/ are mainly unaspirated and sometimes voiced. Thus, cues for voicing for initial stops can be obtained from the timing differences between glottal and supraglottal events (Brown, 1977: 30; Ladefoged & Maddieson, 1996: 50; Lisker & Abramson, 1971: 767; Weismer, 1980: 428).

While the contrast in homorganic stops in English is mainly one of aspiration, Arabic follows a binary system of presence or absence of glottal pulsing during the closure period of the stop (e.g. Flege & Port, 1981: 126; Yeni-Komshian, Caramazza & Preston, 1977: 35). More importantly, while both VOICED stops in English and VOICELESS stops in Arabic have been described as falling in the short lag region, there is a marked difference between the two categories in that English VOICED stops are shorter and seem to occupy the 0-25ms end of the short lag range whereas Arabic VOICELESS stops are longer and tend to be slightly aspirated. In Figure 5.2 I propose a general (though simplified) view of the places English and Arabic stops occupy along the VOT continuum.

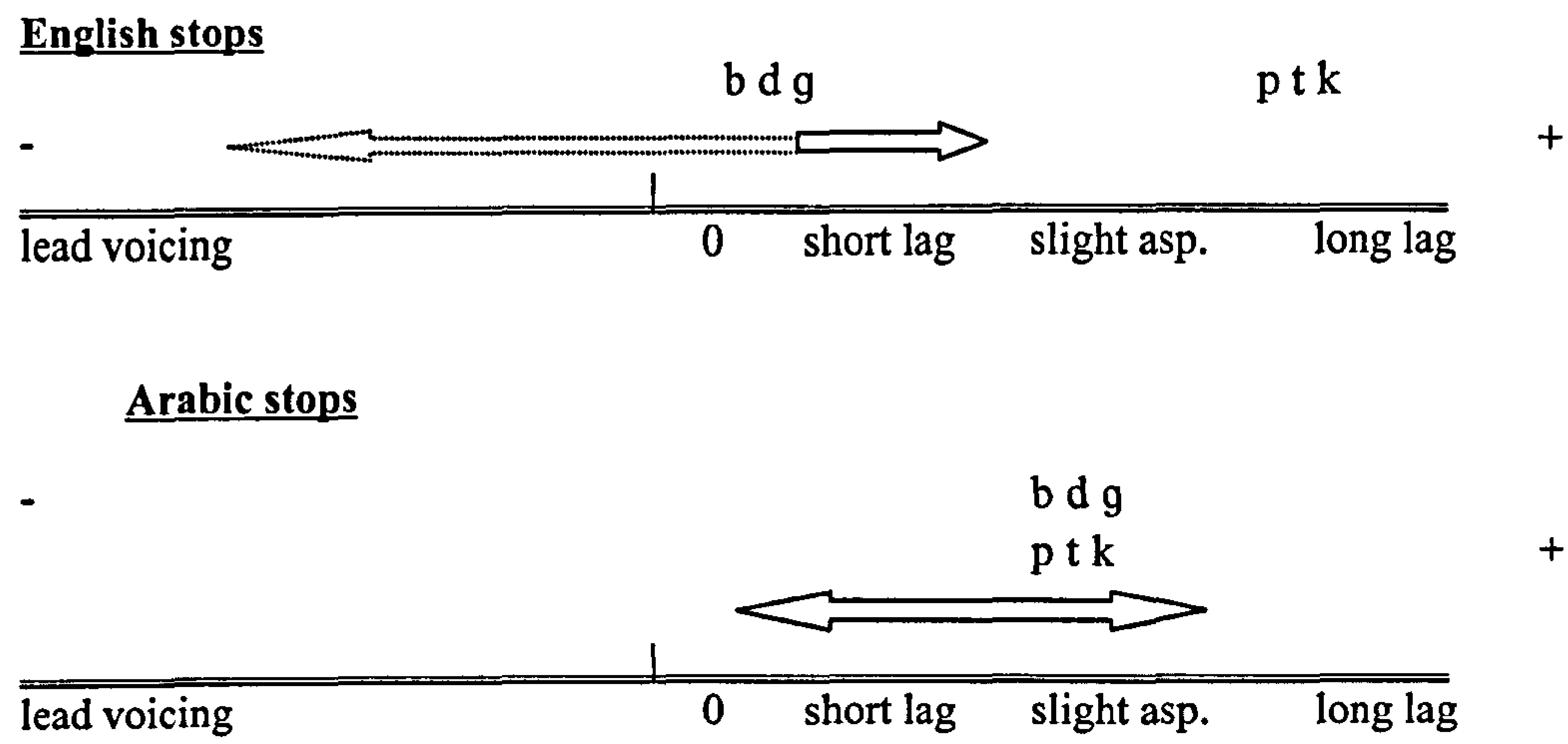


Figure 5.2. Schematic representation of the VOT continuum which shows the relationship between English and Arabic stops. (Adapted from Deuchar & Clark, 1995: 25).

As can be noted from the diagram, while there is overlap between the VOT range for Arabic VOICELESS stops and that of English VOICED stops, there is an important and fine distinction between the two categories, as the positive values for VOICED English stops seem to dominate the left end of the short lag range, while those for Arabic VOICELESS stops dominate the right end and tend to be slightly aspirated in some dialects. Results from my study will be used to examine whether the pattern depicted in Figure 5.2 applies for data from the British English and Lebanese Arabic monolingual and bilingual speakers. The place that English VOICED stops and Arabic VOICELESS stops occupy along the VOT continuum will be particularly important for the examination of whether the bilingual subjects keep the VOT ranges for the two categories separate.

Finally, it is worth noting that the VOT patterns described in Table 5.12 constitute only one aspect of the way voicing patterns in word-initial stops differ in English and Arabic and give us an indication of the complexity of the task faced by a child learning both languages. There are of course other important physiological and acoustic cues that are specific to each language and that play an important role not only in differentiating between VOICED and VOICELESS stops in each language, but also in providing the necessary detail for the native acquisition of the stops. Therefore, VOT patterns in the two languages must not mask the fact that there are other fundamental differences in the production of stops in each language, including their place of articulation (e.g. Nasr, 1966), context- and accent-specific allophonic variation (e.g. the wide range of /t/ variants in certain varieties of English found by Foulkes et al, 1999), and other important acoustic cues such as burst intensity, spectral shape, and formant frequencies in the following vowels (e.g. Mitleb, 1984a; 1984b).

When such differences are considered, it is less likely that sets of stops that are labelled /p t k/ and /b d g/ in the two languages can be seen as phonologically 'similar'. Still, this is the assumption adopted in many cross-linguistic studies, and as a result of that many studies on bilingual acquisition have set out to examine how bilingual children manage to override such 'similarity' and to learn the language-specific phonetic details for each of their languages. This issue will be discussed in the next section.

5.6 VOT in Bilingual studies

Several investigations into the phonological acquisition of bilingual children have compared the production and/or perception of VOT in the subjects' languages, especially where the two languages differ in their use of the VOT continuum, as in English and French (e.g. Caramazza et al, 1973; Cutler, Mehler, Nirris, & Segui, 1989; Elman, Diehl, & Buchwald, 1977; Hazan & Boulakia, 1993; Watson, 1991; 1995), English and Spanish (e.g. Bond, Eddey, & Bermejo, 1980; Deuchar & Clarks, 1995; Flege & Eefting, 1987; Konefal & Fokes, 1981; Yavas, 2002), English and Portuguese (e.g. Sancier & Fowler, 1997; Rocca & Marcelino, 1999), or English and Panjabi (Heselwood, & McChrystal, 2000). In all cases, the languages being examined are described as having similar phonological contrasts between their stops due to the binary presence or absence of phonological voicing (apart from Panjabi which has a three-way contrast), but as differing in their phonetic realisation of the voicing contrast in that English follows the short lag-long lag distinction whereas French, Spanish, and Portuguese follow the voicing lead-short lag distinction. The aim of most studies is to establish whether bilinguals develop separate codes for their languages or whether they use a common code, usually that of the language they have had exposure to the most. Factors that are considered to influence acquisition include country of residence, language of greater exposure, status (political, social, etc.) of each language, age and order of acquisition of each language, and language dominance (degree of bilingualism).

The general consensus is that bilinguals are able to adapt their production mechanisms according to the systems of each language, but that signs of 'interlanguage interference' are inevitable, usually from the strong or dominant language to the weaker. There are, however, differences in opinion with regards to the initial acquisition of the contrast in each language, that is, whether the child starts with, for instance, a set of VOICED and VOICELESS stops that are common for both languages and later acquires the different phonetic implementation rules for each language, or whether the child learns two different sets of VOICED and VOICELESS stops for each language from the start. In Sections 5.6.1 and 5.6.2, I review some of the results from studies on the perception and/or production of bilinguals.

5.6.1 Perception studies

Studies of VOT perception are more numerous than production ones. As with other aspects of bilingual processing, the opinion is divided with respect to whether bilinguals are sensitive to the perceptual cues for the VOICED/VOICELESS models that are appropriate for each of their languages, or whether they use a common (universal) model. Perceptual studies not only test whether VOT is perceived differently by bilinguals according to the language presented, but also whether bilinguals are sensitive to other cues to the contrast that have different perceptual weight in the languages examined.

When testing the use of VOT as a perceptual cue for the voicing contrast, the usual practice is to present bilinguals with synthetic (but occasionally natural) speech-like continua of stops in which VOT values have been manipulated. The stops are then embedded in language-appropriate example words or passages, and the subjects are asked to listen to them and to identify the words. Their categorisation is interpreted in terms of the 50% crossover in their labelling functions and its relevance to the language in question, i.e. the point at which 50% or more of the subjects' responses change from one voicing category to the other. The presence of a phoneme boundary shift in the subjects' perception of categories depending on the language they are listening to is usually taken as evidence for their ability to maintain strict separation between the processing strategies in their two languages, i.e. as evidence of 'code-switching' at the phonemic level.

In some studies, bilinguals have been found to have only one speech segmentation strategy, that of the dominant language (Caramazza et al, 1973; Cutler et al, 1989). In others, bilinguals show language-based difference in categorisation, even if the effect is generally smaller than the one found when comparing them with monolingual subjects from each language (Elman et al, 1977; Hazan & Boulakia, 1993; Slawinski and Wiigs, 1999; Watson, 1995).

Among the early studies that show a significant perceptual shift by bilinguals is that conducted by Elman et al (1977), who tested three groups of adult subjects: monolingual English speakers, monolingual Spanish speakers, and bilingual English-Spanish speakers. Although the subjects' task was to identify nonsense syllables that varied along the voicing dimension from /ba/ to /pa/, natural rather than synthetic stimuli were used, and the test tapes included one- and two-syllable filler words along with the nonsense syllables. The tokens were preceded with language-appropriate instructions to write the 'word' heard. The analysis concentrated on 'ambiguous' syllables, i.e. the ones with a VOT in the short lag region, because the difference in stimulus identification in the two languages was expected to take place around this region. As expected, there was a large difference between the monolingual English and Spanish subjects in the number of /b/

responses to the stimuli, with the English speakers almost always identifying them as /b/ and the Spanish speakers as /p/. More interestingly, bilingual subjects showed varying degrees of identification shift according to the language of instruction depending on their degree of bilingualism. While the more balanced bilinguals showed complete identification shift and a performance that was similar to the monolinguals of each language, less balanced bilinguals showed less significant boundary shifts according to the language set. Still, less balanced bilinguals exhibited monolingual performance in one of the two language conditions, and approached but did not reach monolingual targets in the other condition (Elman et al, 1977: 973).

Similar but more complex results were obtained by Zampini & Green (2001), who tested both production and perception of VOT and closure duration by Spanish-English bilinguals (closure duration plays a significant role in distinguishing between VOICED and VOICELESS stops in Spanish but not English). The subjects showed English and Spanish monolingual-like production abilities on both acoustic measures, but their perceptual abilities were monolingual-like only with respect to VOT. The subjects showed sensitivity to closure duration of the stop in both English and Spanish mode, which the authors interpreted as a residual attention that is always present (in analogy to slight activation of the bilinguals' other language even when they are in a monolingual mode).

Another study that shows a significant perceptual shift by bilinguals is one that was conducted by Slawinski and Wiigs (1999), who examined differences in the categorical perceptions and phoneme boundary locations of bilingual speakers of Polish and English. The study is original in that it compares speech perception and categorisation in monolingual English speakers and English-Polish bilingual speakers on both English and Polish. Ten subjects from each group listened to a continuum of [bi-pi] synthesised syllables varying from –40 to +50ms VOT values in Polish and English contexts and were asked to identify the syllables. They also rated the quality and the intelligibility of each syllable on a three-choice rating scale. While monolingual speakers used categorisation cues for VOICED and VOICELESS stops that are based solely on their native language regardless of the perceptual set, bilingual listeners showed significantly different categorisations of VOICED and VOICELESS stop consonants depending on the language context in which they were asked to make their choices. Moreover, the syllable ratings of the two groups differed depending on what counts as 'good quality' [p] or [b] in each language based on VOT values. Individual differences between bilingual subjects were also observed mainly as a function of the age of acquisition of their second language, which ranged between six and 12.

Watson (1995) conducted a perceptual experiment to determine whether English-French bilingual children's development of the categories underlying the voicing contrast resembled that of monolingual control groups. He investigated two groups of bilinguals (one residing in England and the other in France) aged six, eight and ten, and compared their performance on a series of perceptual tests to that of monolingual subjects from each language. The experiments consisted of two synthetic VOT continua whereby VOT was manipulated in the first one and F1 in the second. The continua were produced with carrier sentences in English and French and were played to the subjects who then indicated their responses by ticking the appropriate box under pictures that illustrated the target words. The major result from the study was that all bilingual groups at all ages responded differently to the VOT continua in the English and French conditions, i.e. there was a lower crossover point in French than in English, representing a different phoneme boundary. However, there were also significant differences between bilinguals and monolingual controls with respect to their categorisation of the VOT continuum, although the results varied depending on the age group. Moreover, the language spoken in the country of residence had the greater influence, with higher category boundaries observed in London-based bilinguals than in Paris-based ones. As for the results from the F1 manipulation, no clear patterns emerged for any of the bilingual or monolingual subjects, although older English children tended to be more capable of responding to it than other groups. Watson concluded that the bilingual child may have two systems, but these may differ in some way from those of the monolingual.

Another study of English-French bilinguals is that by Hazan & Boulakia (1993), who examined the production and perception of two groups of bilinguals aged between 15 and 43 and living in Great Britain and France in order to account for the bias due to the language of immersion. Monolinguals from each language were also taped in their country of residence and served as controls. The authors were interested in finding out whether bilingual listeners use separate voiced-voiceless prototypes in both languages, and whether factors such as language dominance and age of acquisition of the second language affect sensitivity to certain acoustic cues in any way. Perceptual tests consisted of tokens of /pɛn/ and /bɛn/ that were produced in French and English contexts and whereby formant frequencies and duration of formant transitions were manipulated in order to evaluate the perceptual effect of spectral cues within the post-consonantal vowels. VOT values ranged between –40 and +40ms, chosen to cover acceptable values in both languages. Production tests consisted of reading carrier sentences and minimal pairs that contained target consonants from each language. The subjects were classified into different groups depending on their strength of bilingualism, which was determined from questionnaire data and from phonetically-trained listener judgments.

Results from production tests showed evidence of code-switching between the two languages in the production of both bilingual groups. The only significant differences between bilinguals and monolinguals were in the production of English /b/ and /p/ by French-dominant bilinguals. Moreover, while there were no significant differences between the productions of early and late acquirers of the L2, strength of bilingualism seemed to play a role in the production of monolingual-like phones.

Results from perceptual tests showed significant difference between the two groups of bilinguals with respect to their labelling of the continua. The phoneme boundary was at a shorter VOT duration for the English-dominant than for the French-dominant bilinguals, which the authors found surprising, as they expected a greater proportion of voiced responses for English-dominant listeners. Similarly to the production results, strong (more balanced) bilinguals were more likely to code-switch than weak (less balanced) bilinguals, and there were significant differences in the labelling behaviour of the two groups. As for the use of perceptual cues, there were clear differences between the French-dominant and English-dominant groups, which were parallel to the differences between the monolingual groups.

While VOT was a dominant cue for French-dominant subjects (and monolingual French subjects) in ambiguous VOT regions (at phonemic boundaries where transition cues conflicted with temporal ones), English-dominant (and monolingual English) subjects showed greater sensitivity towards cues in the vowel onset. Such results agree with those obtained by Caramazza et al (1973) on Canadian French, reviewed below. Still, Hazan & Boulakia emphasised the importance of individual differences in cue weighting for the English monolingual and bilingual subjects, whereas the French monolingual and bilingual ones had more homogeneous behaviour. For the majority of bilinguals their dominant language rather than the language of presentation affected their sensitivity to perceptual cues, which provides support for the theory that, even in balanced bilinguals, one language does dominate for certain aspects of language processing (Cutler et al, 1989).

Yet another study of English-French bilinguals is that by Caramazza et al (1973), who examined the perception and production of Canadian English and French bilinguals and monolinguals. As opposed to the other studies that examined the role of VOT in French, the authors in this study reported that VOT was not a sufficient cue, neither at the perception nor at the production level, to distinguish homorganic stops in the Canadian French dialect that they examined. The study was conducted with 20 bilinguals and 10 monolinguals from each language, all aged between 17 and 25. The subjects read stop-initial French and English words that were later spectrographically analysed, and then labelled synthetic speech sounds from 'stop + vowel' syllables that differed in VOT only. The bilinguals were all native speakers of French who had learned their English by no later than their 7th birthday and who were judged as proficient in both languages following a series of assessments.

In the perceptual experiment, the authors found that VOT served as a perceptual cue when labelling homorganic stops for the monolingual English but not French subjects (the rate of change from one category to the next was fast and monotonous for the English group but not for the French one). As for the bilingual subjects, their 50% crossover values occupied intermediate positions relative to those of the monolingual groups, and so did the rate of change from one category to the next and the amount of variability. As opposed to the monolingual French subjects, the bilinguals did appear to use VOT as a perceptual cue, but their behaviour was similar on sets of data from both languages, which suggests that their perceptual decisions followed the same criteria for both English and French.

In the production experiment, VOT once again proved to be an important variable for voicing distinction in English but not in French (overlap in phonemic categories occurred only in French). The bilingual subjects produced voicing results that were clearly different for the two languages and that were better than their perceptual behaviour, but their values were more aligned with the French monolinguals in the French mode than the English monolinguals in the English mode. On the one hand, their VOT values for the VOICELESS consonants yielded significant differences from those of the monolingual English subjects. On the other hand, their production of VOICED stops showed no appreciable difference in the two languages, and ranged from the voicing lead to short lag region. The authors concluded that VOT control was important for the subjects only at phonemic boundary regions and relatively unimportant at other points in the productive range where the information carried by VOT is phonemically irrelevant. The overall results were explained in terms of language interference. Since the subjects acquired English as their second language, the authors concluded that their phonological system was subject to interference, in this case a unidirectional one from the stronger language to the weaker.

Summary of perception studies

While most of the studies reviewed in this section have demonstrated that VOT perception is influenced by a bilingual experience, there are a few cautionary remarks that need to be taken into consideration. First, very few studies have examined the perceptual abilities of infant bilinguals (e.g. Bosch & Sebastián-Gallés, 2002) in order to test whether the bilingual initially perceives one or two sets of VOT categories for their languages. We know from studies on infant perceptual abilities (Chapter One, Section 1.5.1) that bilingual and monolingual infants can perceptually discriminate different linguistic systems at birth, and later concentrate on the contrasts of their ambient language(s). Therefore, in the case of bilinguals exposed to two languages from birth, there is no reason to believe that they will not be able to perceive two different VOT systems for their languages.

The studies that were reviewed in this section, however, concentrate on later stages of development and describe bilingual second language acquirers. Consequently, factors such as age and order of exposure to each language, language of greater exposure, majority language in the country of residence, and language dominance all appear to play a role in the uneven perceptual abilities in each language that are reported by the researchers. But one needs to remember that the performance of a child on a perceptual experiment at any time during his/her development will represent an intermediate- rather than a final state of his/her perceptual abilities. Not only do children experience a gradual development of adult-like perceptual abilities, but there is also evidence that adult-like consistency might not be complete until the second decade of life, even in monolinguals (Hazan & Barrett, 2000). Moreover, the experience of a bilingual upbringing can lead to the use of different strategies from those used by monolinguals. This may be due to a difference in processing patterns between those who are already bilingual at the time of their first phonological development, and those who are consecutive bilinguals. Though most bilinguals show a significant perceptual shift between their two languages, variability in their responses to perceptual cues often depends on their degree of bilingualism, which necessitates the importance of paying attention to individual results along with more general patterns.

Variability in responses to perceptual cues, however, has been shown to be a feature of monolingual as well as bilingual speech, due to the differences in the nature of cues and the weight given to them among languages. Beyond the age of one, research shows that the rate of development of categorical perception varies across languages. In a study conducted by Simon & Fourcin (1978) on the perceptual abilities of monolingual English and French children aged 2 to 14, two major findings were reported. First, both English and French children progressively learned to use acoustic patterns that are relevant for distinguishing stops in their languages as the basis of categorical labelling. However, while English-speaking children tended to acquire the ability to respond in a categorical fashion to a VOT continuum by the age of four, French children did not acquire categorical labelling until the age of eight or nine. Secondly, while English-speaking children gradually became aware of the perceptual salience of variations in the onset frequency of F1 following the stops, the F1 cue played no role in VOICED-VOICELESS perception for French.

The authors concluded that the French children's task might be difficult due to the fact that they might be concentrating on other features that are normally found in the production of voicing oppositions in adults. For instance, VOICELESS bursts will have more high-frequency energy than VOICED bursts, especially in stressed initial position. For this reason, judging the children's perceptual abilities by manipulating F1 for example may be too complex a perceptual task for young children. This issue was also discussed in Section 5.5.2 with respect to child production abilities (Scobbie, Gibbon, Hardcastle & Fletcher, 2000) and similar conclusions were drawn. Future perceptual tests will have to reconsider whether bilinguals use different acoustic cues than monolinguals altogether or maybe attribute different weightings to perceptual cues.

Another issue which has rarely been discussed in studies of VOT perception (and production) in bilinguals concerns variability in the input that the children receive and that might affect their responses to cues for the voicing contrast. For instance, although both Hazan & Boulakia (1993) and Caramazza et al (1973) examined VOT perception in French-English bilinguals, dialect-specific differences with respect to perceptual cue weighting in the French varieties that were examined (Parisian *versus* Canadian) proved significant for the interpretation of the bilinguals' behaviour. Moreover, within the same dialect, we know that there is abundant inter- and intra-speaker variability in production in general, and consequently in the realisation of the stop voicing contrast. It is therefore difficult to assume simple targets for the children to perceive, and some of the 'interference' in the perception of language-specific phonemic boundaries that was reported for bilingual subjects might actually reflect variable input.

Finally, results from perceptual studies should be interpreted with care due to the problems associated with using synthetic stimuli. One such problem is that the stimuli may be insufficiently natural to evoke responses typical of normal speech behaviour Thomas (2000). Another is that the stimuli are usually manipulated for one or two parameters that are being studied (e.g. VOT, F1 transitions, etc...), but may lack other acoustic cues that differ in languages and that are important cues for the perception and/or production of stops in some languages/accents and for certain subjects (e.g. Simon & Fourcin, 1978). The bilingual being tested will have to rely heavily on the cues available in the stimuli and may not be able to alter his/her stimulus identification according to the language of presentation. This issue has been raised in monolingual experiments as well. Thomas (2000: 20) notes that labeling of stimuli in an experiment is not the same as perception in conversations. When subjects are asked to identify stimuli as one of two or three choices, they probably focus on particular cues and thus may modify their means of perception from what they use in normal conversations.

5.6.2 Production studies

As in perception studies, the main debate in VOT production concerns the issue of whether the bilingual child develops one VOT system for both languages or one for each. But unlike perception which can be tested at a very early age, production cannot of course start to be tested until the beginning of speech in the child. Therefore, children may well conceive of VOICED-VOICELESS cognates as two distinct phonological categories in each language long before they can actually produce them in their speech (e.g. Scobbie, Gibbon, Hardcastle & Fletcher, 2000).

The normal development pattern for VOT production is for all stops to be produced in the short lag region initially, with production extended, as appropriate for the language, to the long lag or prevoiced ranges (e.g. Macken & Barton, 1979, 1980; Zlatin & Koenigsknecht, 1976). For some children, this process may not be complete until the age of four. The following is a review of some of the studies carried out on VOT production in bilinguals.

Deuchar & Clark's (1995) study (also reported in Deuchar & Quay, 2000) is one of the few reports on the early bilingual acquisition of the voicing contrast. The authors conducted a case study of Deuchar's daughter's acquisition of English and Spanish from birth, and collected data at the ages of 1;7, 1;11, and 2;3. Data were recorded in the form of daily diary records at home and studio recordings during which the child was asked to name objects in pictures designed to elicit target words in both languages. At age 1;7, all the child's productions in both languages fell within or not far outside the short lag range (0-20ms for labial and alveolar stops, and 0-40ms for velars), but there were not enough tokens to draw conclusions about the establishment of a voicing contrast at this age. At age 1;11, a contrasting pattern started to emerge in the two languages. In the English data, the means of the VOICELESS stops were all longer than their VOICED counterparts (though only significantly for /t d/), and were similar to the adult pattern. In Spanish, however, the pattern did not apply to all places of articulation and there were no tokens with voicing lead as expected for VOICED stops. Clearer indications of a voicing contrast appeared at the age 2;3 in both English and Spanish. The differences between the means for VOICED and VOICELESS stops were greater than before, though only significant for English, and the pattern of short lag *versus* long lag was found in both languages. Statistical tests yielded significant differences between English VOICED and VOICELESS stops, English VOICELESS and Spanish VOICED stops, and English VOICELESS and Spanish VOICELESS stops. No significant difference, however, was found between Spanish VOICED and VOICELESS stops, and Spanish VOICED and English VOICED stops, while the difference between Spanish VOICELESS and English VOICED stops narrowly missed significance. Table 5.13 shows the mean VOT values obtained in English and Spanish at age 2;3.

Table 5.13: Mean VOT measurements (ms) for word-initial stops produced by the Spanish-English bilingual child in Deuchar & Clark's study at age 2;3.

	p	t	k	b	d	g
English	62	76	78	17	22	30
Spanish	37	37	42	26	29	28

While the English data conform to adult English patterns by showing long lag for /p t k/ and short lag for /b d g/, the Spanish ones do not conform to the adult Spanish patterns, though there are beginnings of a VOT distinction based on contrasting lag measurements. The authors suggested a progress between ages 1;11 and 2;3 and some indication of a distinct voicing contrast beginning to be established in Spanish, but that does not involve an adult-like lead versus lag difference. Analysis of the parents' speech in Spanish was also undertaken by Deuchar & Clark (1995), since the parents were the only source of Spanish to the child. The father, who is a native speaker of Spanish, produced all his VOICED stops with voicing lead and his VOICELESS ones with short lag. The mother (Deuchar), who is a native speaker of English and who leaned Spanish in adulthood, produced all but three of her VOICED stops in the short lag region and her measurements were similar to those of the child at age 2;3.

The authors conclude that there was a progression from a lack of system in either language at age 1;11 (though there were indications of a system beginning to be established in English) to the establishment of a clear voicing system in English at age 2;3, and the beginning of a distinct system in Spanish (Deuchar & Quay, 2000: 45). Such results underline the importance of examining adult input before attempting to interpret the bilingual child's behaviour. However, it is interesting to note that the findings are similar to those of monolingual English and Spanish children (Macken & Barton, 1979; 1980) and monolingual French children (Allen, 1985). In all those reports, there seems to be a developmental constraint on the production of voicing lead by children under the age of four due to vocal tract dynamics, while short and long lag follow adult patterns from an early age.

The acquisition of voicing lead in bilinguals does seem to develop at a later stage. Data on the production of older Spanish-English bilinguals are available from two earlier studies by Bond et al (1980) and Konefal & Fokes (1981). Bond et al (1980) measured the production of two consecutive bilingual sisters aged between four and seven, the latter being severely language-disordered. The subjects had Spanish as their mother tongue, and had started acquiring English at school. The younger child was reported to speak English and Spanish fluently and had clearly differentiated VOT patterns for her stops in the two languages, with values corresponding to monolingual norms. She prevoiced her Spanish VOICED stops, aspirated her English VOICELESS stops, and produced her VOICELESS Spanish stops and VOICED English stops in the short lag region. The older child, who was still at the two-word stage due to her language disorder, produced all her stops in the short lag region, a pattern typically found for younger children.

Similar results were found by Konefal & Fokes (1981), who measured the production of three sisters aged four, seven, and ten who also had Spanish as their mother tongue, and the oldest sister was language-disordered. The two normal children showed clearly differentiated patterns for their two languages, but the youngest child produced her VOICED stops in the short lag region, with the VOICELESS stops spanning the upper part of that region and the long lag range. However, the lack of prevoicing in her VOICED Spanish stops is not necessarily ascribable to English-Spanish interference, but is possibly part of her normal developmental pattern. Evidence for the latter option can be seen from the VOT patterns for the seven-year-old, which correspond to monolingual norms of both languages. As for the oldest, disordered child, she showed considerable overlap between the categories in both languages.

Despite the similarity between monolinguals and bilinguals in the eventual development of language-specific VOT patterns, there is evidence that their realisation of the voicing contrasts do have some significant differences. Watson (1991) compared bilingual and monolingual development of a number of acoustic features associated with the voicing contrast in English and French, including VOT and F1 frequency. The subjects were two groups of bilinguals and monolinguals aged six and ten and adult controls from each language. One striking result from the study was the variability in the production of all subjects, including the monolinguals' parents. For instance, in French, VOICELESS stops were sometimes aspirated while VOICED ones were produced in the short lag region. Both tendencies were found in children more than adults. Two disparities were found between bilinguals and monolinguals. First, bilingual subjects produced the VOICELESS series with more aspiration than the monolinguals in both age groups and in both languages. Second, bilingual subjects showed a significant distinction between their VOICED and VOICELESS stops in both languages not only with respect to VOT, but also with respect to the onset frequency of the first formant in following vowels, which is normally salient in English but not in French (Simon & Fourcin, 1978). Despite those two differences, Watson maintains that the overall productions of the monolinguals are not distinguishable from those of the bilinguals (impressionistic judgements were made by the subjects' teachers and other native speakers who listened to the tapes). It seems that bilinguals can use different production routines from those of the monolinguals without this being perceptible to other native speakers.

Success in doing so at later stages, however, depends on the bilingual's experience with the two languages and several factors surrounding the acquisition process. Rocca & Marcelino (1999) investigated English and Portuguese VOT production in five adult bilinguals who had different linguistic backgrounds. The subjects had different degrees of exposure to and social identification with the relevant communities, as well as in the amount of phonetic training they had received in each language. They ranged from being intermediate speakers of English as a foreign language to being proficient bilingual speakers with one of the two languages as the first language. The experiment consisted of reading Portuguese and English words in carrier sentences whereby the target words had one of the VOICELESS stops /p t k/ in word-initial and word-medial positions. VOT results ranged from partial phonetic approximation in English by the least experienced native Portuguese speaker to complete command of the native-like VOT patterns in each language by the early bilingual speaker. While the native Portuguese speaker who was an intermediate English learner transferred the short lag VOT of Portuguese onto his English production of /p t k/, the other subjects showed varying degrees of keeping the short lag/long lag distinction separate in their production of English and Portuguese VOICELESS stops that correlated with their proficiency levels.

Amongst the production studies reviewed in this section, Grosjean & Miller (1994), Heselwood & McChrystal (2000) and Yavas (2002) are some of the few studies that I am aware of that take into account the language mode of the bilingual. Grosjean & Miller (1994) measured VOT in the onset of code-switches in the production of French-English bilingual adults with little, if any, foreign accent in either language. The subjects were asked to retell stories in French, in French with English code-switches, and in English. The stories involved a number of character names that could be said in English and in French and that started with /p/, /t/, and /k/ (Paul, Tom, Carl). Results showed that, while the bilinguals showed a significant difference between English and French VOT values, the English code-switched values were significantly different from the French values and similar to the English values. The results suggested that, in bilingual speech production, there is no phonetic momentum of the base language that carries over to the guest language (at least when the bilinguals master the phonetics of the two languages). Switching from one language to another appears to involve a total change not only at the lexical but also at the phonetic level (Grosjean, 2000: 455).

Heselwood & McChrystal's (2000) study is unique in terms of the type of voicing contrast that was compared (three-way contrast for Panjabi *versus* binary for English), as well as the sociolinguistic interpretation that the authors attempted to establish by incorporating the VOT patterns of their ten-year-old Panjabi-English subjects with results from English accent features examined in these Bradford-born subjects. English and Panjabi stops differ on both phonological and phonetic terms. With regards to the phonological difference, there are only two sets of stops in English (/b d g/ and /p t k/), but three in Panjabi (/b d g/, / p t k/ and /p^h t^h k^h/). As for phonetic differences, English VOICED stops are more likely to be produced with short lag while Panjabi ones are normally fully voiced (Heselwood & McChrystal, 1999). Moreover, English VOICELESS stops are aspirated, while Panjabi has two sets of VOICELESS stops, one produced with short lag and the other with aspiration.

The authors taped 19 ten-year-old English-Panjabi bilinguals in English-only and Panjabi-only sessions, and compared their English production with that of monolingual controls and their Panjabi production with that of older generations of Panjabi speakers living in Bradford, some of whom were bilingual whereas others had very little knowledge of English (Heselwood, & McChrystal, 1999). While ten-year-old the subjects produced English VOICELESS stops with a similar amount of aspiration to that of the monolingual controls, their English VOICED stops had a significantly higher amount of prevoicing than that of the monolinguals and was interpreted in terms of influence from Panjabi. Moreover, the Panjabi VOICED stops were either produced with prevoicing or with short lag, the latter option causing the three categories of Panjabi stops to be collapsed into two and becoming a prevailing trend in Panjabi-English bilinguals under 25 growing up in Bradford (Heselwood & McChrystal, 1999). More importantly, the authors noted that the amount of prevoicing in English and Panjabi varied a great deal depending on the subjects, with some of them producing prevoicing consistently in both languages whereas others producing no prevoicing in either language. Such results drove the authors to conclude that both languages must be activated in the bilinguals regardless of the language session, as the influence between the two languages seems to be mutual.

Yavas (2002) examined /p t k/ production in ten seven-year-old Spanish-English bilinguals who had Spanish as their L1 and lived in South Florida. A Spanish-English bilingual teacher asked the subjects to repeat sets of English sentences and mixed English-Spanish and Spanish-English sentences with target English and Spanish stops. However, Yavas (2002: 345) combined the VOT values of the three sets of sentences when presenting the results, and justified doing so by noting that there was no significant difference between the VOT of the English stops produced in monolingual and mixed sentences. One could argue that, although the sentences varied from monolingual to mixed, it does not necessarily follow that the children's languages were more or less activated during any of the productions, as in all cases they were interacting with the same bilingual teacher, and therefore the interlocutor might have played a role in keeping both their languages activated. When analysing the combined results, Yavas (2002: 345-346) noted mutual influence from both languages on the bilinguals' productions in that some of the Spanish VOT measures were 'brought up' for some subjects due to influence from English, while some of the English VOT measures were 'brought down' for some subjects due to Spanish influence. The criteria used for the English and Spanish norms, however, were taken from Lisker & Abramson's (1964) VOT means for English and Spanish stops produced by adults, and therefore these means did not necessarily serve as a suitable basis upon which the children's values were judged as lower or higher. In all cases, the author found that regardless of the language influence, the children had VOT values for their stops in each language that differed in the right direction (VOT for Spanish was always shorter than for English) despite the fact that the difference was not significant for all subjects and for all places of articulation.

Summary of production studies

Overall, production studies suggest that bilinguals behave in ways that are at once distinct from monolinguals but also very similar to them. During the early stages of production, it is often difficult to look for signs of differentiation between the VOT systems of two languages since (i) the little amount of data produced by the child at this stage does not allow for any firm conclusions to be drawn (Deuchar & Clark, 1995) and (ii) the production of certain phonetic categories (e.g. voicing lead) is at the mercy of articulatory maturation and therefore affects both bilingual and monolingual production (e.g. Macken & Barton, 1979; 1980). At a later stage, some studies have found that bilinguals do develop monolingual-like VOT production patterns for each language (e.g. Bond et al, 1980), whereas others note that there might be still be subtle differences that are imperceptible to the listener (e.g. Watson, 1991). In the case late and/or less proficient bilinguals, on the other hand, noticeable signs of interference from the VOT patterns of the first and/or dominant language have been documented (Rocca & Marcelino, 1999).

While there have been ample discussions of factors that influence bilingual speakers' VOT production such as age, language dominance, and country of residence, two issues still need further discussion and investigation. The first one concerns the input that the children receive in both languages. On one hand, Deuchar & Clark (1995) showed how the child's adoption of a given VOT pattern (in this case voicing lag instead of voicing lead for Spanish VOICED stops) might actually be traced to the input that she receives (in this case that of her native-English mother who also spoke Spanish to the child). On the other hand, Watson (1991) showed that both bilingual and monolingual adults exhibit a lot of variability in their productions, which might in turn be displayed in the children's productions. Not all studies have taken adult input into consideration.

The second factor concerns the language and social context from which the bilinguals' utterances were extracted and analysed. For instance, was the bilingual communicating with a monolingual or with a bilingual interviewer? In the latter case, were two languages used simultaneously or did the researcher follow a one-language-per session method? Perception studies reviewed in Section 5.6.1 have shown that bilingual are capable of 'switching' their VOT phonemic categories depending on the language of instruction. Similarly, in Chapter One, we discussed the effect of the interlocutor on the bilingual's linguistic choices in terms of deciding whether to use one language or two during a conversation, and, in the second case, how much code-switching to engage in. Similar decisions might apply on a more subtle and detailed level, for instance, with regards to the VOT patterns produced for a given language depending on the interlocutor and on whether one or both languages are activated. So far, the few studies available disagree on whether or not language mode affects the VOT patterns that are produced for a given language (Grosjean & Miller, 1994; Heselwood & McChrystal, 1999; Yavas, 2002). In all three studies the subjects belong to sizable bilingual communities and often find themselves in a bilingual mode. It would be interesting to investigate cases of individual bilingualism where the bilinguals frequently find themselves switching between monolingual and bilingual language modes.

5.6.3 English-Arabic studies

There are no studies on VOT in English-Arabic bilingual subjects, but a series of studies have examined VOT in the production of English as a foreign language by Arabic speakers (Flege, 1980, Flege & Port, 1981; Fokes, Bond, & Steinberg, 1985; Port & Mitleb, 1983). These will be reviewed below.

Flege (1980) and Flege & Port (1981) conducted several experiments aimed at examining certain acoustic correlates in the production of Arabic and English stops by groups of Saudi Arabians learning English as a foreign language. Results for the Arabic experiment were discussed in Section 5.5.3, and showed that word-initial VOICELESS stops in Saudi Arabic are slightly aspirated, while VOICED stops are produced with almost categorical voicing lead. In the English experiment, analysis was made of the production of /p t k/ and /b d g/ by a group of six Americans and two groups of six Saudi Arabians who differed according to their length of residence in the US (eight months *versus* 39 months). The subjects were asked to produce English CVC words in the carrier sentence 'I say _ again to Bob', and measurements were made of VOT, vowel duration, and stop closure duration in initial and final stops in the target words.

Results showed significant differences between the productions of the American subjects and those of the two Saudi groups, due to 'non-English phonetic characteristics' in the English produced by the Saudi Arabians (Flege & Port, 1981: 133). VOT values that were produced by the Saudi subjects in English were similar to Arabic VOT values that were found for Saudi speakers in the Arabic experiment. Moreover, the Saudi subjects made the closure intervals of VOICELESS stops longer than those of VOICED stops in word-initial position, a contrast that was significant in all but one case and that was not found in the production of the American subjects.

Flege & Port (1981) suggested that the Saudi speakers were 'carrying over' Arabic phonetic features of the stop voicing contrast onto their English stop production. Further evidence for this claim was found in the patterns for vowel and stop closure duration in word-final position. There was no significant difference in VOT production between the two Saudi groups, although the group that had spent a longer time in the US had slightly higher VOT values for VOICELESS stops (Flege & Port, 1981: 135). The only difference between the two Saudi groups was in the duration of word-final stops, whereby the more experienced group produced different consonant duration for VOICED and VOICELESS stops (similar to the English pattern),[13] although such difference in duration contrast is not found in Arabic. The authors concluded that the more experienced Saudis were approximating phonetic norms of English, although the performance varied greatly between- and within- individuals, depending on the voicing correlate examined.

More importantly, the absence of /p/ from Arabic did not prevent any Saudi subjects from showing their awareness of the phonetic differences between /p-b/ in English by generalising the duration difference between /t-d/ and /k-g/ found in Arabic to the English /p-b/ contrast and exaggerating it at times. However, their laryngeal control differed during their production of /p/ as compared to /t/ and /k/, due to the fact that a substantial number of their /p/ tokens was produced with glottal pulsing during the closure period. Flege & Port concluded that it may be more difficult to learn to control a new pattern of glottal-supraglottal timing than one involving purely supraglottal timing. Still, the authors maintained that the behaviour of second language learners shared similarities with child first-language acquisition, mainly with respect to the exaggeration

[13] Closure durations of stops in final position in English are subject to a large voicing effect (VOICELESS greater than VOICED) in both stressed and unstressed positions (cf. Stathapoulos & Weismer, 1983)

of certain phonetic dimensions found in the adult target and the gradual approximation of new phonetic dimensions.

Port & Mitleb (1983) later replicated certain aspects of the experiment by Flege & Port (1981), using a group of American subjects and two groups of Jordanian subjects. The Jordanian speakers differed in that one group had lived in the US for 12-16 months whereas the second had never lived in an English speaking country before. All subjects read test CVC words embedded in the carrier sentence 'He tried to say __ again', where the consonants examined were the minimal pairs /p-b/ and /t-d/. Measurements included VOT for VOICELESS stops, vowel duration, final consonant closure duration, and sentence duration.

VOT results showed that while the American subjects produced long lag for /p/ and /t/, the two Jordanian groups produced short lag values that were almost half the duration of the American VOT values. Although /p/ is not part of the Jordanian Arabic inventory, the subjects still managed to produce it with short lag, although some of the tokens had weak voicing in the closure period. Like in Flege & Port's (1981) study, the two Jordanian groups differed only in the final consonant closure duration, whereby only the more experienced Jordanians managed to produce longer duration for /p/ than for /b/, and shorter duration (less than 30-40ms) for /t/ and /d/ (evidence for flapping).

Port & Mitleb (1983: 228) concluded that adult foreign language learners can learn new phonological rules, as evidenced by the Jordanians' success in producing a novel /p-b/ contrast in English, but that it is more difficult for them to change temporal implementation rules, as evidenced by the Jordanians' use of a short lag for their /p-t/ production (similar to the Arabic pattern).

Among the studies reviewed in this section, that by Fokes, Bond, & Steinberg (1985) is the only one that was conducted with Arabic children rather than adults. The study had various aims, including investigating whether young Arab children who are learning English as a second language can acquire VOT patterns that are appropriate for initial English stops and whether acquisition rate increased in line with the increase of age of exposure to English. The authors taped 12 children from a variety of Arabic backgrounds who had arrived in the US at different ages and who had varying degrees of exposure to English. The subjects were divided into two groups on the basis of age (24 to 60 months and 84 to 135 months), and were later regrouped according to experience with English (two to 12 months *versus* 18 to 54 months). Table 5.14 shows the mean VOT values obtained from each of the four groups of children.

Table 5.14: Mean VOT measurements (ms) for word-initial English stops produced by the Arab children in Fokes, Bond, & Steinberg (1985: 85)

	p	t	b	d
Younger	84	88	-7	6
Older	77	71	2	16
Less experienced	70	71	2	16
More experienced	92	102	-10	9

All the children seemed to have acquired long lag for VOICELESS stops despite considerable variability in the subjects' performance as shown in individual results presented later in the study. Moreover, most of the children (apart from two less experienced speakers) used similar VOT distinction for apical and bilabials stops, even though the voicing contrast is lacking for bilabials in Arabic. As for VOICED stops, although the means in Table 5.14 showed a tendency to be produced with short lag, individual results also showed that several subjects actually used prevoicing (Fokes, Bond, & Steinberg 1985: 86-87). However, this issue was not commented on by the authors, who concentrated on the fact that the children were using a great deal of phonetic variability that is similar to that of young native English children. As there was no significant difference in VOT production between any of the groups, whether compared by age or experience, the authors concluded that the children had acquired the English voicing contrast regardless of how little exposure they had had to English, and that children learning English are therefore more successful in resembling native speakers than their adult counterparts. Though the study provides us with valuable data on Arab children learning English, a lot of issues are left unaddressed, such as how the VOT values that these children produced in English compared with their VOT production in Arabic, and the role that the different dialectal backgrounds of the children might have played in influencing their production in English (the subjects' countries of origin included Saudi Arabia, Libya, Palestine, Kuwait, and Sudan).

5.7 Aims of the current study

There are few instrumental studies with data on the phonetic/phonological basis of VOT patterns in Arabic in general, and in the Lebanese dialect specifically. As mentioned in the review of Arabic studies in Section 5.5.3, the only data available on VOT production in Lebanese adults come from the study conducted by Yeni-Komshian et al (1977), although the authors used reading material from Classical Arabic rather than the Lebanese dialect. Apart from this study, no research has been done on the VOT production of monolingual or bilingual Lebanese adults or children. This chapter is therefore designed to investigate the following questions:

1 Do English-Arabic bilinguals acquire separate VOT patterns for each of their languages?
2 Are their patterns of production in each language similar to those of the monolingual controls in the study?
3 Are the patterns for the monolingual subjects in this study similar to the ones normally described in the literature and therefore expected for each language?
4 Are there signs of influence from one language on the other in the bilinguals' production and what are the factors that affect such influence?

5.8 Procedure

5.8.1 The material

Data from all 23 speakers in this study were used for this experiment. It was hypothesised that including monolinguals would help refine the targets available for the bilinguals by examining whether the stark contrast between the two languages that has been suggested in the literature with respect to VOT production also stands for Lebanese Arabic. The aim was also to obtain a sample of child VOT values in each language using the same data elicitation techniques and analysis methods as the ones used with the bilinguals. By eliciting naturalistic Arabic data in the subjects' dialect, the study differs from most previous ones that have used reading material from MSA or Classical Arabic.

All the English and Arabic words that have initial stops in prevocalic position and that were produced in isolation by the children during the picture-naming activities and by the parents during the reading list activities were extracted from the tapes. The tokens chosen for analysis conformed to the following criteria:

- English and Arabic words were chosen with similar environments following the stops (in order to control for the intrinsic differences mentioned in Section 5.4). Attempts were made to have a balanced number of tokens that fell either in a stressed or in an unstressed position, that were either mono- or polysyllabic, and had comparable post-consonantal vowels (though this last condition was difficult to implement due to the different vowels the subjects produced in English and in Arabic).
- For each language, only the two sets of /b d g/ and /p t k/ stops were analysed. As mentioned earlier, /p/ and /g/ are not part of the sound inventory of Lebanese Arabic, but do occur in loan words. Arabic /tˤ dˤ/ were not analysed as there were very few occurrences of these stops in the children's speech.

5.8.2 Analysis

All the target words produced by the subjects were digitised onto a PC running the Sensimetrics SpeechStation 2 software. Wideband spectrograms and waveforms were generated of the words. A total of 2020 tokens were analysed for all children and adults. First, auditory analysis was made for each token to determine the identity of the stop and whether it was heard as VOICED or VOICELESS. Then, using instrumental analysis, binary judgement of the presence or absence of glottal pulsing during stop closure was made and noted for each of the target stops. For voiceless stops, VOT was measured from the beginning of the release burst to the start of periodicity in the following vowel. For prevoiced stops, VOT was measured from the onset of periodicity in the closure period to the start of the release burst. This almost always corresponded with the start of periodicity in the following vowel. Measurements were all taken from the waveforms, though visual inspection of the corresponding spectrograms was also used as back up in cases where background noise in the recordings caused disturbance to the waveform.

Mean VOT and standard deviations were calculated for the consonants in each language and for each subject. The results were then compared and interpreted in terms of the subjects' sociolinguistic background, mainly their age, language input, use, and dominance. As two of the bilingual children were the subjects of a previous pilot study in which measures of their VOT production were also taken (Khattab, 1998), their results from this study were compared with the former, which allowed me to track their development over a 18-month period.

5.9 Results

5.9.1 Adults: English

Results for the adults are discussed first in order to examine the type of VOT patterns that are likely to be available in the children's environment. Figures 5.3, 5.4, 5.5 and 5.6 show mean VOT measures and VOT distribution (in ms) for the monolinguals' parents and the bilinguals' parents first presented in two groups (Figure 5.3) and then individually (Figures 5.4-5.6).

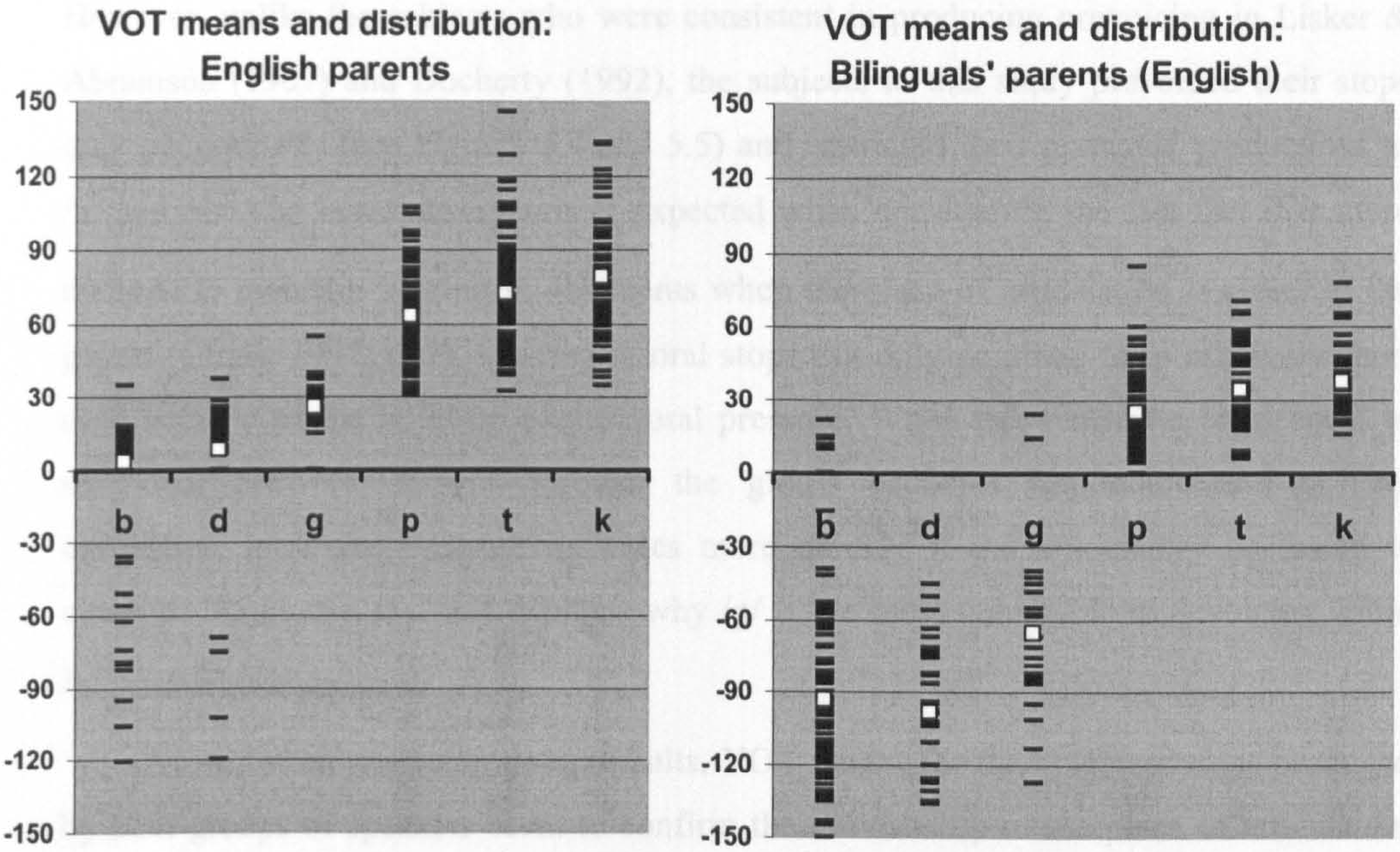


Figure 5.3: Mean VOT values (white squares) and distribution (in ms) for the monolinguals' parents (left) and the bilinguals' parents (right). N = 822.

On the whole, the mean values obtained for the adults in this study follow the expected pattern for each group of speakers. While the monolinguals' parents mainly produce short lag with some voicing lead for their VOICED stops and long lag for their VOICELESS stops, the bilinguals' parents apply the Arabic VOT patterns on their VOT production by mainly producing voicing lead for their VOICED stops and short lag to slight aspiration for their VOICELESS stops. But there are interesting observations that need to made.

First, although there is overlap between the VOT values for VOICED and VOICELESS stops produced by both groups of speakers, individual results (Figures 5.4 and 5.5) show that most speakers do keep the distributions for each pair of stops quite separate. Second, although there is overlap between the VOT distributions for /p t k/ by the bilinguals' parents and /b d g/ by the monolinguals' parents, there is a significant difference between the two groups of distributions (t-tests significant at $p << 0.01$ for all three places of articulation). The findings suggest that the two distributions belong to different phonetic categories, one of slight aspiration for /p t k/ by the bilinguals' parents, and the other of short lag for /b d g/ by the monolinguals' parents.

Third, a number of the VOICED stops that were produced by the monolinguals' parents had voicing lead, which, along with results found by Lisker & Abramson (1967) and Docherty (1992), confirms the fact that descriptions of the phonetic realisations of English VOICED stops should not be restricted to one single category (voicing lag).

However, unlike the subjects who were consistent in producing prevoicing in Lisker & Abramson (1967) and Docherty (1992), the subjects in this study prevoiced their stops only occasionally (see Figures 5.4 and 5.5) and restricted their phonated productions to /b/ and /d/. The latter observation is expected when considering the fact that it is more difficult to maintain voicing in obstruents when the place of articulation is closer to the glottis (Ohala, 1997: 687). Voicing in oral stops can only continue for a relatively short time because of the build-up of intra-oral pressure. When this reaches a level equal to subglottal pressure, airflow through the glottis becomes impossible and voicing terminates. Intra-oral pressure increases more quickly if the articulatory occlusion is closer to the glottis, and this explains why /g/ is the most vulnerable to devoicing while /b/ is most resistant to it.

Third, when analysing group results, VOT means for the voiceless stops produced by both groups of speakers seem to confirm the universality of the place of articulation effect in that VOT seems to increase as the place of articulation for the stop moves further back in the mouth. However, when looking at individual results (Figure 5.4), this study offers further support to observations made by Docherty (1992) about exceptions to this apparently universal rule and the importance of looking at individual differences. While Docherty (1992: 130) found a slight tendency for alveolars to have longer VOT than velars, in this study, EM5 has a slightly higher VOT mean for /p/ (60.90ms) than for /k/ (58.50ms), while EF10 has a VOT mean for /p/ (61.06ms) that is much higher than that for /t/ (53.94ms) and actually closer to that for /k/ (63.46ms).

Next, the VOT patterns found for /p/ and /g/ as produced by the bilinguals' parents confirm the fact that they have acquired their production despite the rare occurrence of these two sounds in their native language, but have applied the Arabic phonetic implementation for VOICELESS sounds by producing /p/ with short lag to slight aspiration and for VOICED sounds by producing /g/ with voicing lead. Similar results have been found for English /p/ produced by Saudi speakers (Flege, 1980; Flege & Port, 1981; and Port & Mitleb, 1983) and the authors suggested that their subjects were showing their awareness of the phonetic differences between /p-b/ in English by generalising the duration difference between /t-d/ and /k-g/ found in Arabic to the English /p-b/ contrast.

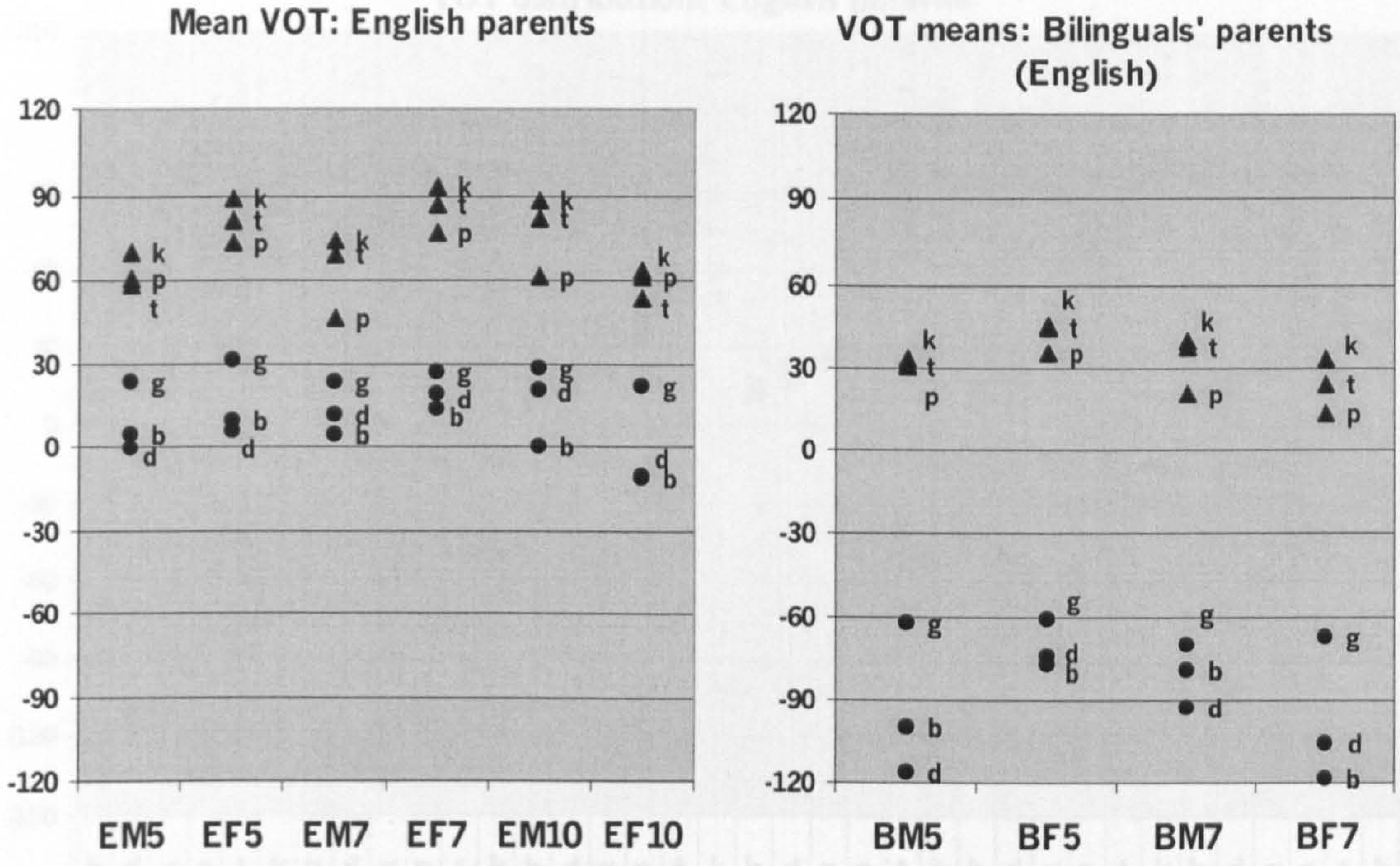


Figure 5.4: Mean VOT values (in ms) for each of the monolinguals' parents (left) and bilinguals' parents (right).

Looking at the VOT means (Figure 5.4) and distributions (Figures 5.5 and 5.6) for each of the individual English and bilingual parents, one can see that there is very little overlap in VOT between VOICED and VOICELESS stops for all of the speakers. Overlap mainly took place when some of the /g/ tokens were produced with slight aspiration (e.g. EF5, EM7, and EM10) so that they had similar and sometimes slightly higher values than some of the VOT values for /p/ (Figure 5.5). On the other hand, some of the /b/ and /d/ tokens in English were produced with voicing lead, but were restricted to 1-3 tokens for most speakers apart from EF10 who produced eight out of 38 VOICED tokens with voicing lead. Moreover, four out of the total of 16 prevoiced tokens produced by all the English speakers exhibited a termination of voicing before the end of the closure period of the stop (Figure 5.7).

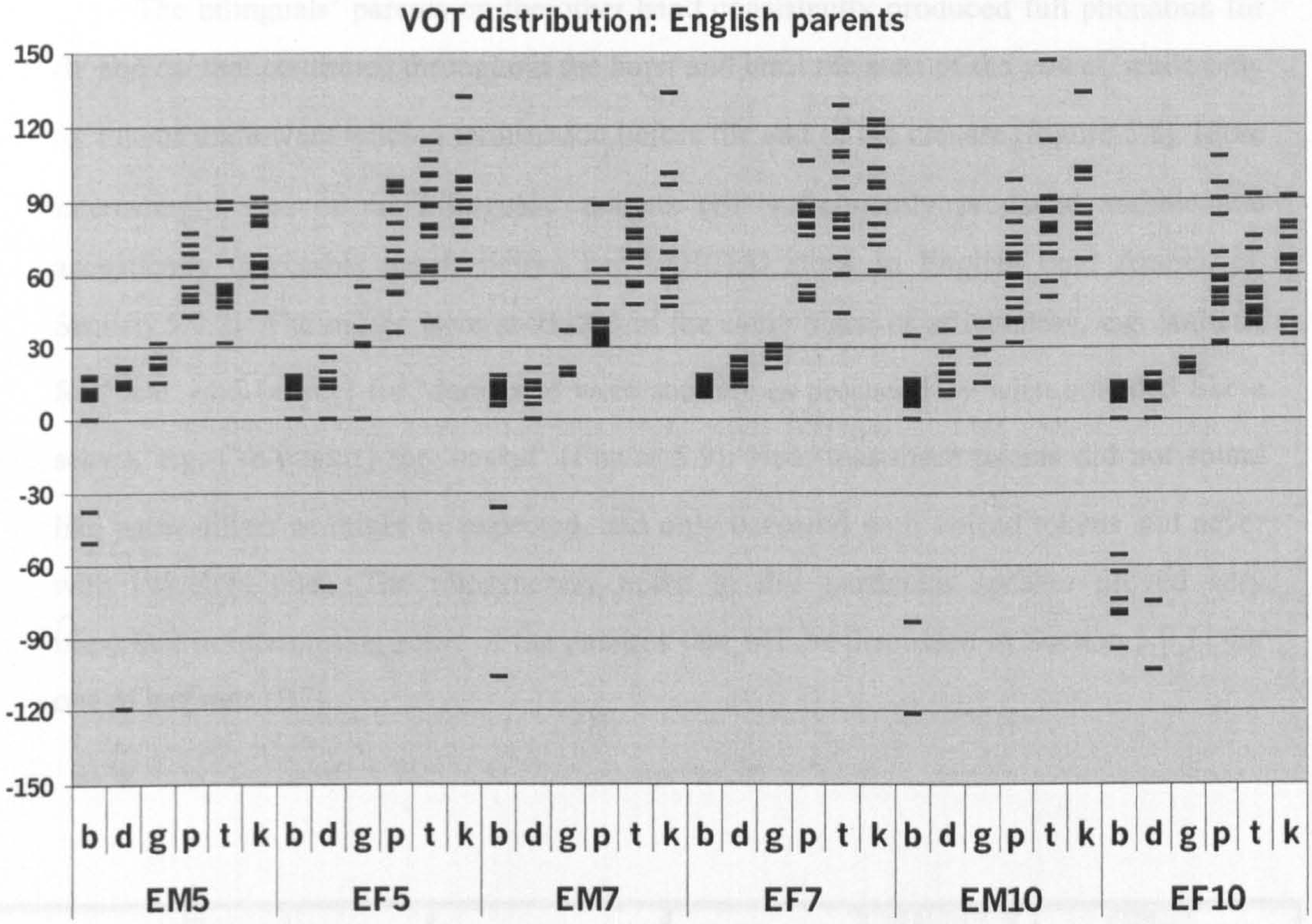


Figure 5.5: VOT distribution (in ms) for each of the monolinguals' parents. N = 482.

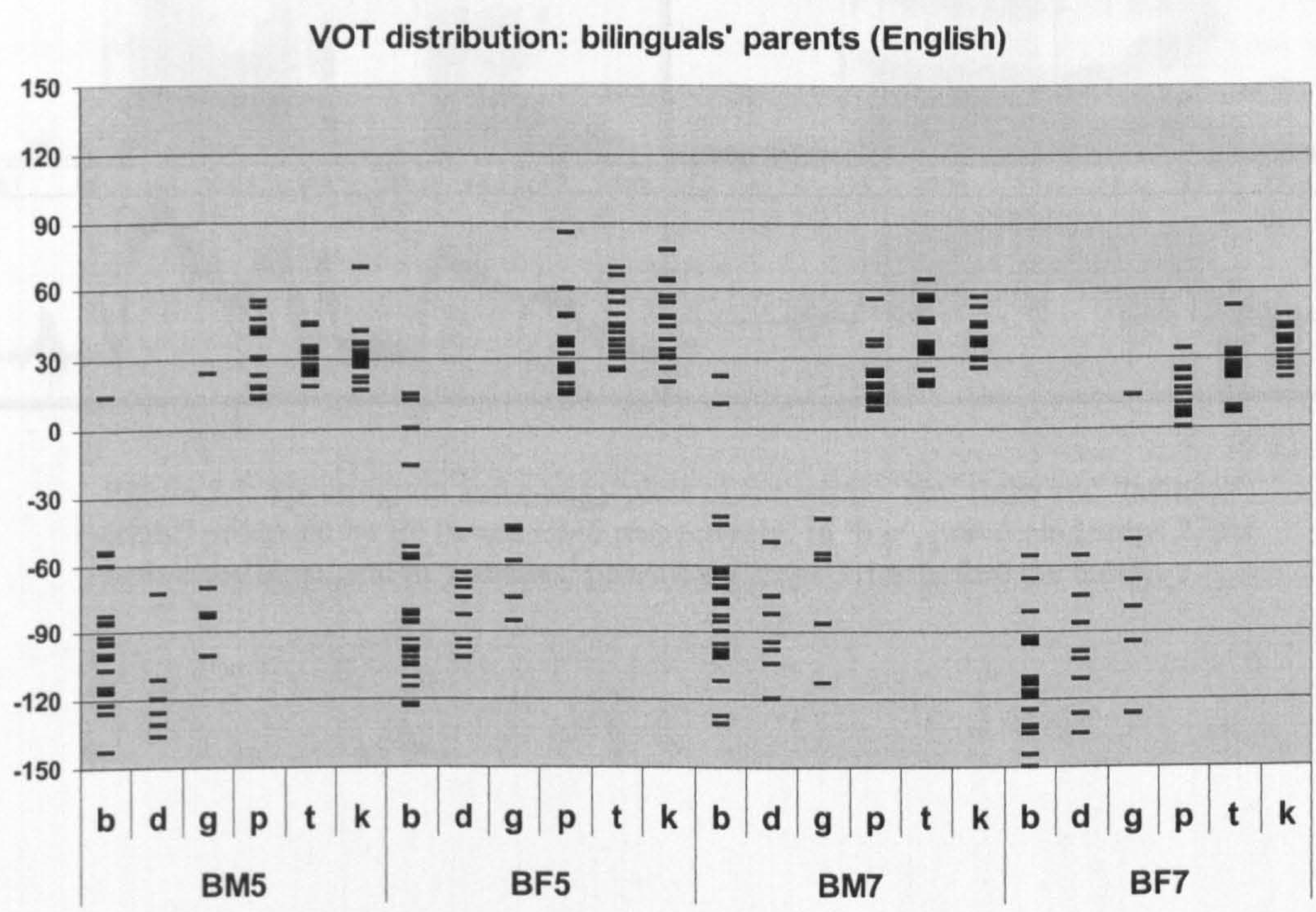


Figure 5.6: VOT distribution (in ms) for each of the bilinguals' parents in English. N = 340.

The bilinguals' parents on the other hand consistently produced full phonation for /b/ and /d/ that continued throughout the burst and until the start of the vowel, while only /g/ tokens underwent voicing termination before the end of the closure (Figure 5.8). More interestingly, one of the bilinguals' parents (BF7) frequently produced audible and acoustically detectable nasals before her VOICED stops in English (and Arabic, cf. Section 5.9.2). The nasals were produced in the same place of articulation, e.g. [mbɛːɾ] for 'bear' and [ndoːɾ] for 'door' and were sometimes preceded by what sounded like a schwa, e.g. [ᵊmˈbaskɪt] for 'basket' (Figure 5.9). Note that these tokens did not sound like pause-fillers as might be expected, and only occurred with voiced tokens and never with voiceless ones. The idiosyncrasy noted in this particular speaker proved very important in interpreting some of the patterns that will be discussed in Section 5.9.11 for one of her sons (B7).

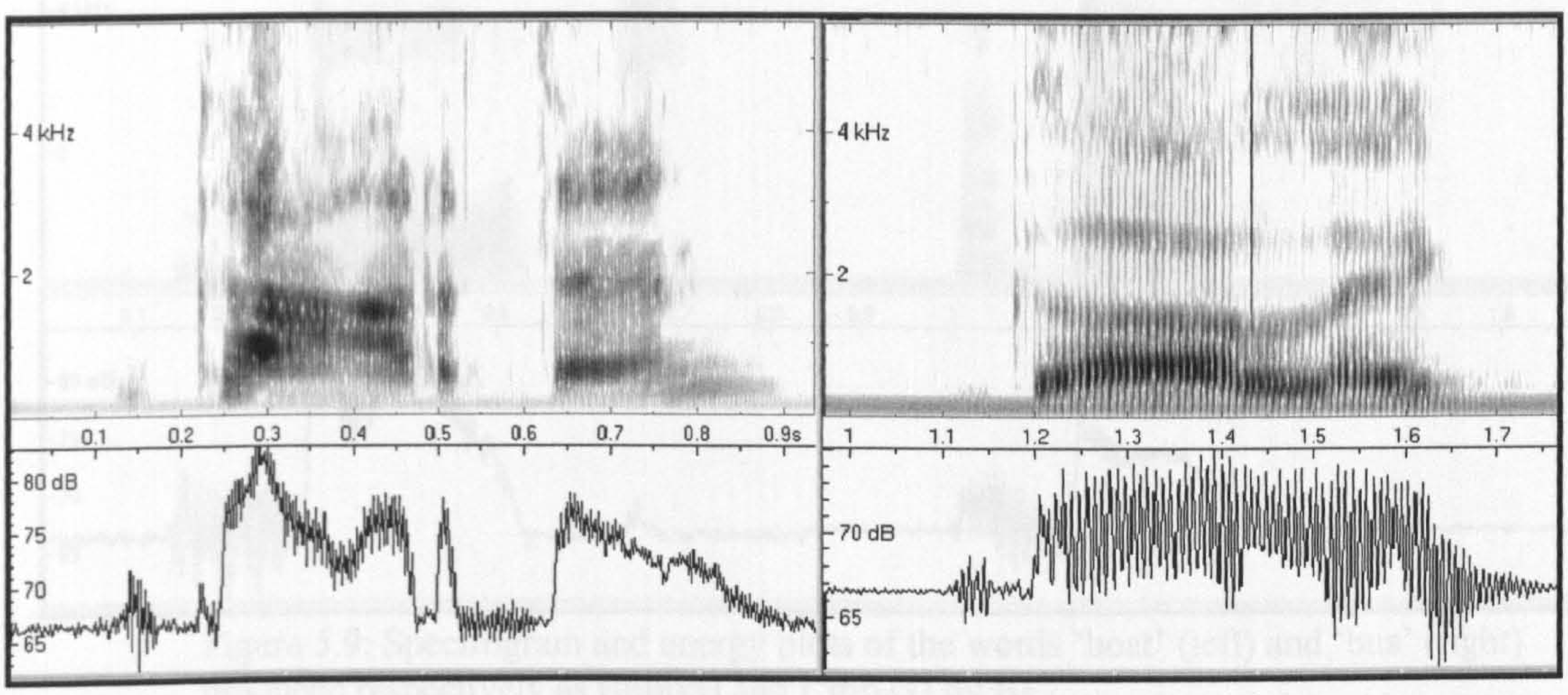


Figure 5.7: Spectrogram and energy plots of the words 'bin' (left) and 'bananas' (right) produced by EF10 and EM5 respectively. In 'bin', prevoicing stops 27ms before the burst, and in 'bananas' prevoicing stops 31ms before the burst.

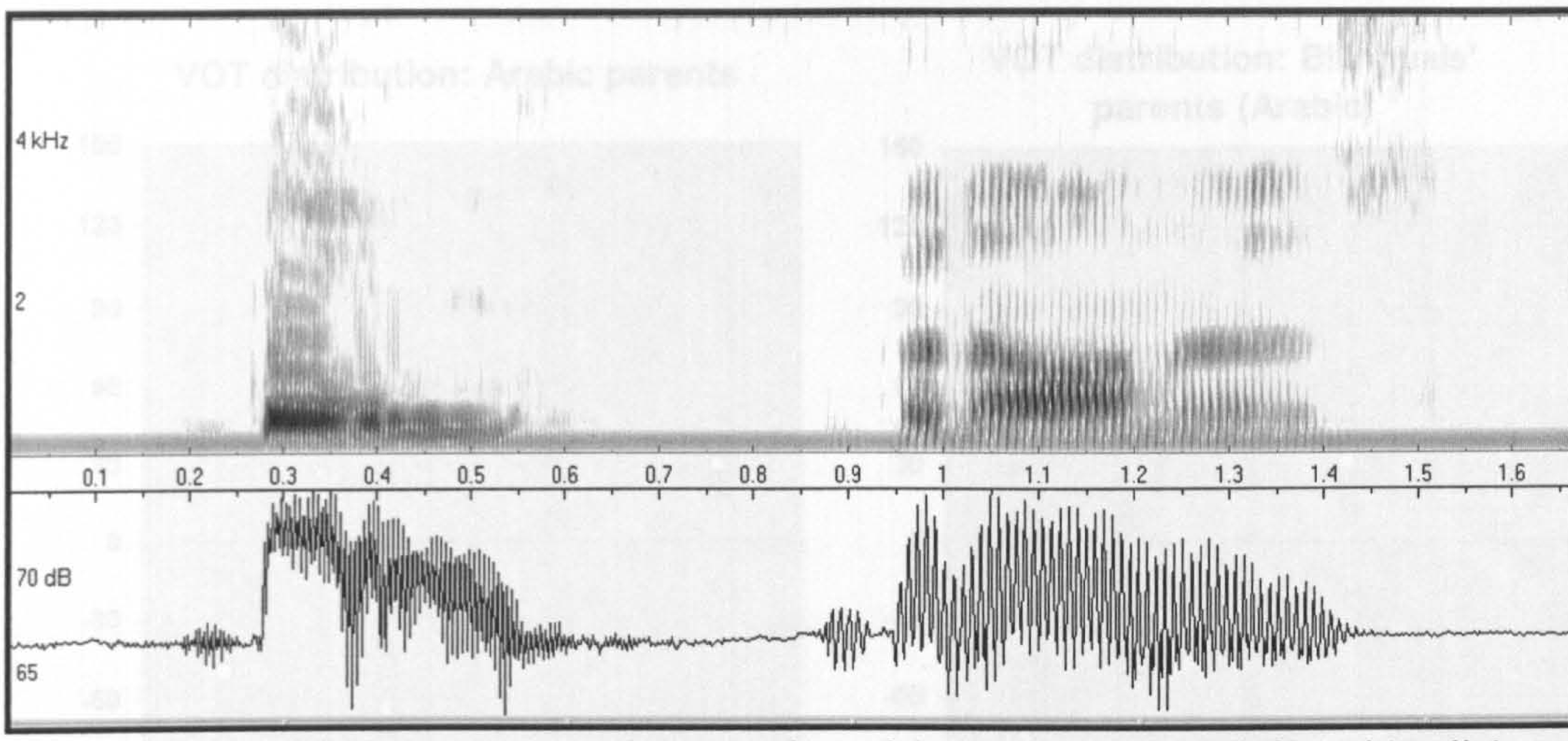


Figure 5.8: Spectrogram and energy plots of the words 'garden' (left) and 'garlic' (right) produced by BF7 and BM7 respectively. In 'garden', prevoicing stops 47ms before the burst, and in 'garlic' prevoicing stops 29ms before the burst.

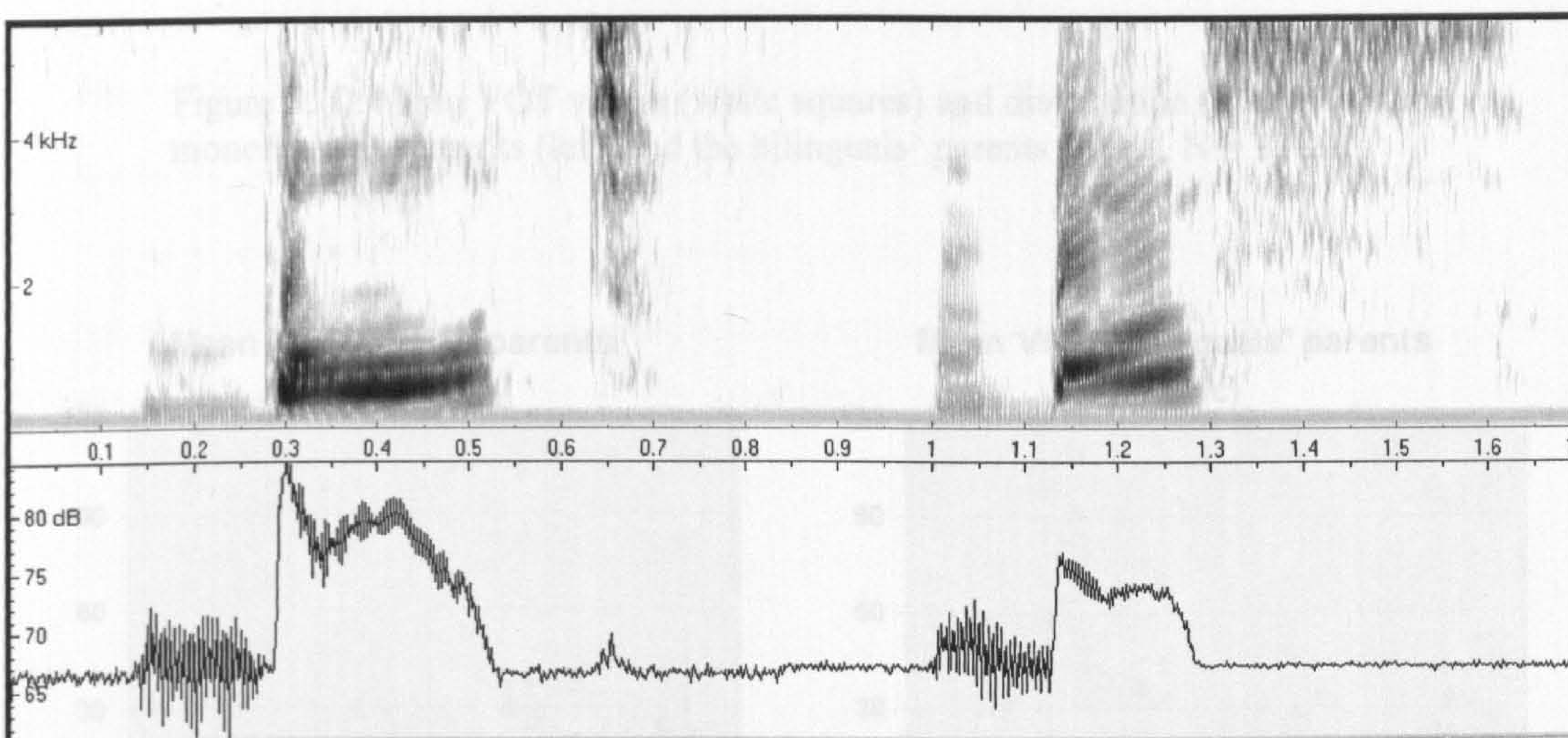


Figure 5.9: Spectrogram and energy plots of the words 'boat' (left) and 'bus' (right) produced respectively as [mbo:t] and [ᵊmbʌs] by BF7.

5.9.2 Adults: Arabic

Figures 5.10-5.13 show mean VOT measures and VOT distribution (in ms) for the monolinguals' parents and the bilinguals' parents first presented in two groups (Figure 5.10) and then individually (Figures 5.11-5.13).

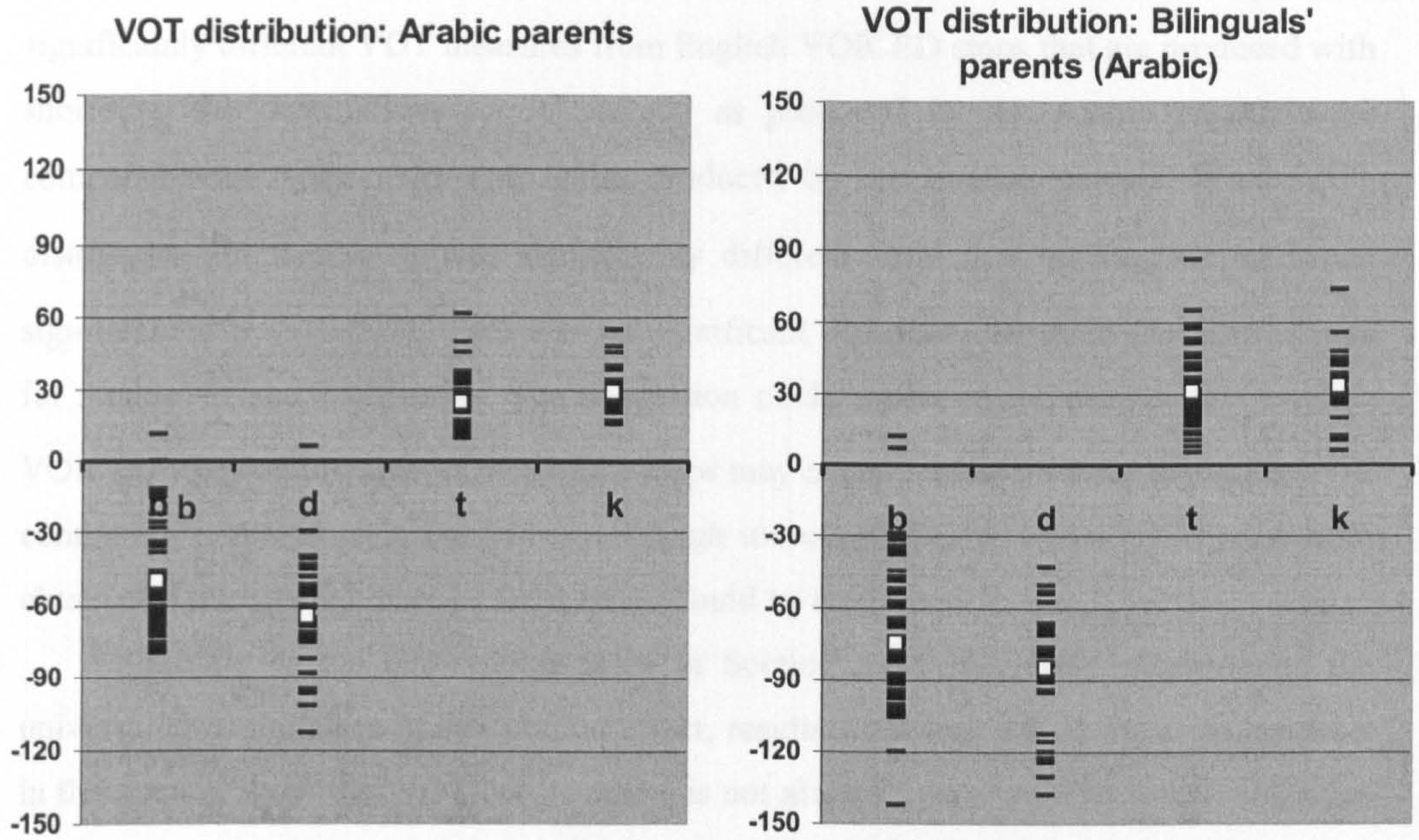


Figure 5.10: Mean VOT values (white squares) and distribution (in ms) for the monolinguals' parents (left) and the bilinguals' parents (right). N = 338

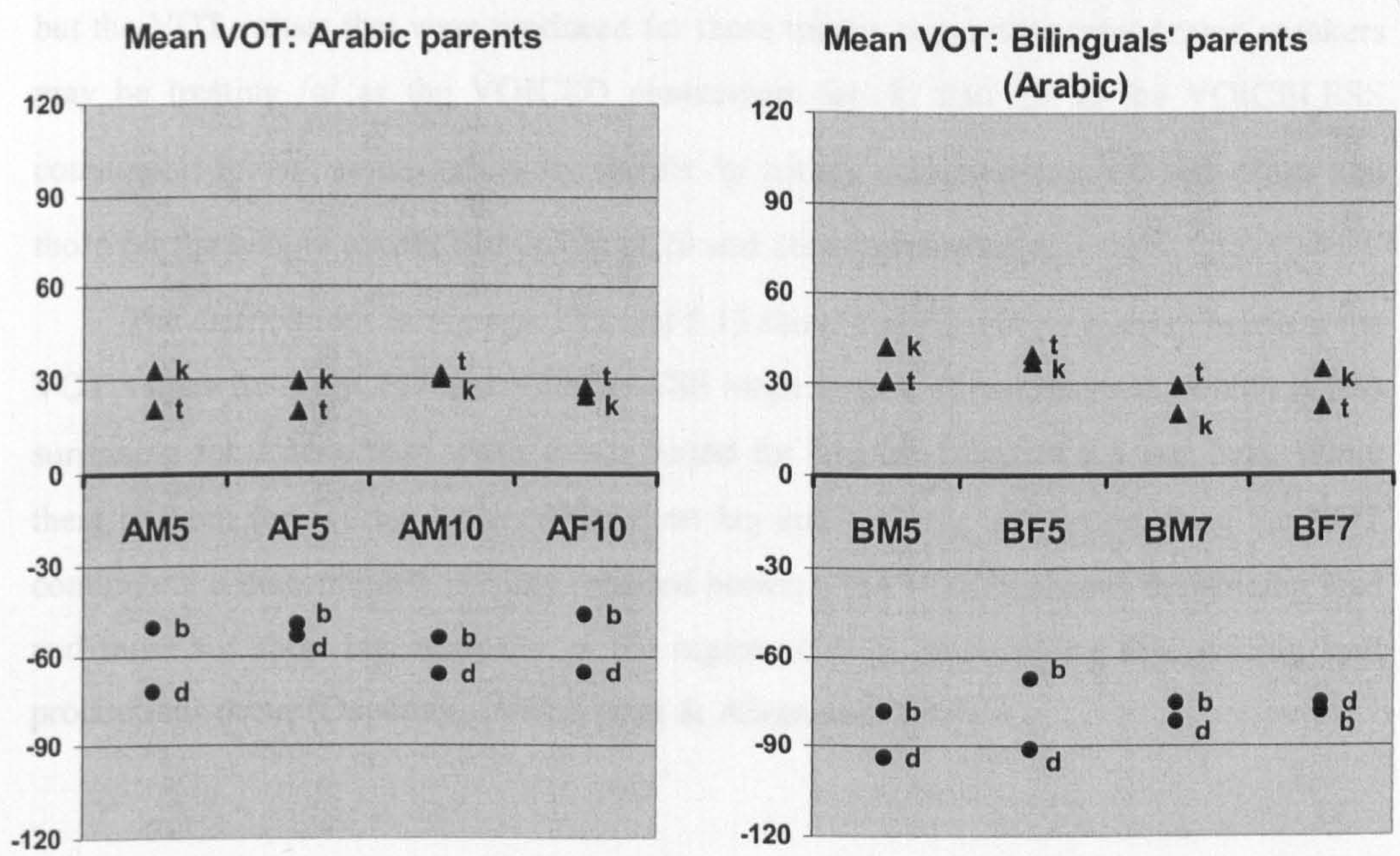


Figure 5.11: Mean VOT values (in ms) for each of the monolinguals' parents (left) and bilinguals' parents (right).

Results from the Lebanese adults in this study are similar to those found for other Arabic dialects (Section 5.5.3) and confirm the fact that the overwhelming VOT pattern for Arabic is that of voicing lead for VOICED stops and short lag to slight aspiration for

VOICELESS stops. In order to examine whether Arabic VOICELESS stops have significantly different VOT measures from English VOICED stops that are produced with short lag, the distributions for /t/ and /k/ as produced by the Arabic parents were compared with those of /d/ and /g/ as produced by the English parents. While VOT distribution for Arabic /t/ was significantly different from that for English /d/ (t-test significant at $p \ll 0.001$), there was no significant difference between the distributions for Arabic /k/ and English /g/. The suggestion made earlier in the chapter that English VOICED stops and Arabic VOICELESS stops may occupy different areas along the VOT continuum is therefore inconclusive, although more significant results could have been obtained if the bilabial place of articulation could be compared.

Similarly to the observation made in Section 5.9.1 about the exception to the universality of the place of articulation effect, results from four out of the eight speakers in this section show that VOT before velars is not always longer than before alveolars, as is evident in the patterns for AM10, AF10, BM5, and BM7 (Figure 5.11). Apart from the results for /b/, /d/, /t/, and /k/, most of the monolinguals' parents and bilinguals' parents produced one or two tokens of /g/ and/or /p/ in loan words like 'pyjama' and 'garçon' (*waiter*). Altogether, there were only six /g/ tokens and two /p/ tokens in the Arabic data, but the VOT values that were produced for these tokens suggest that the Arabic speakers may be treating /g/ as the VOICED counterpart for /k/ and /p/ as the VOICELESS counterpart for /b/, as the values for the six /g/ tokens were between -60 and -97ms and those for the two /p/ tokens had values of 20 and 26ms respectively.

The distributions in Figures 5.12 and 5.13 show that there is no overlap between the VOT values for VOICED and VOICELESS stops in any of the speakers, which is less surprising for Arabic than when it was found for English (Figures 5.5 and 5.6). While there is room for overlap between the short lag and long lag categories along the VOT continuum, a discontinuity is often reported between the VOT measures for voicing lead and those for short lag, normally in the region of 0 to 30ms where few voicing lead productions occur (Docherty, 1992; Lisker & Abramson, 1967).

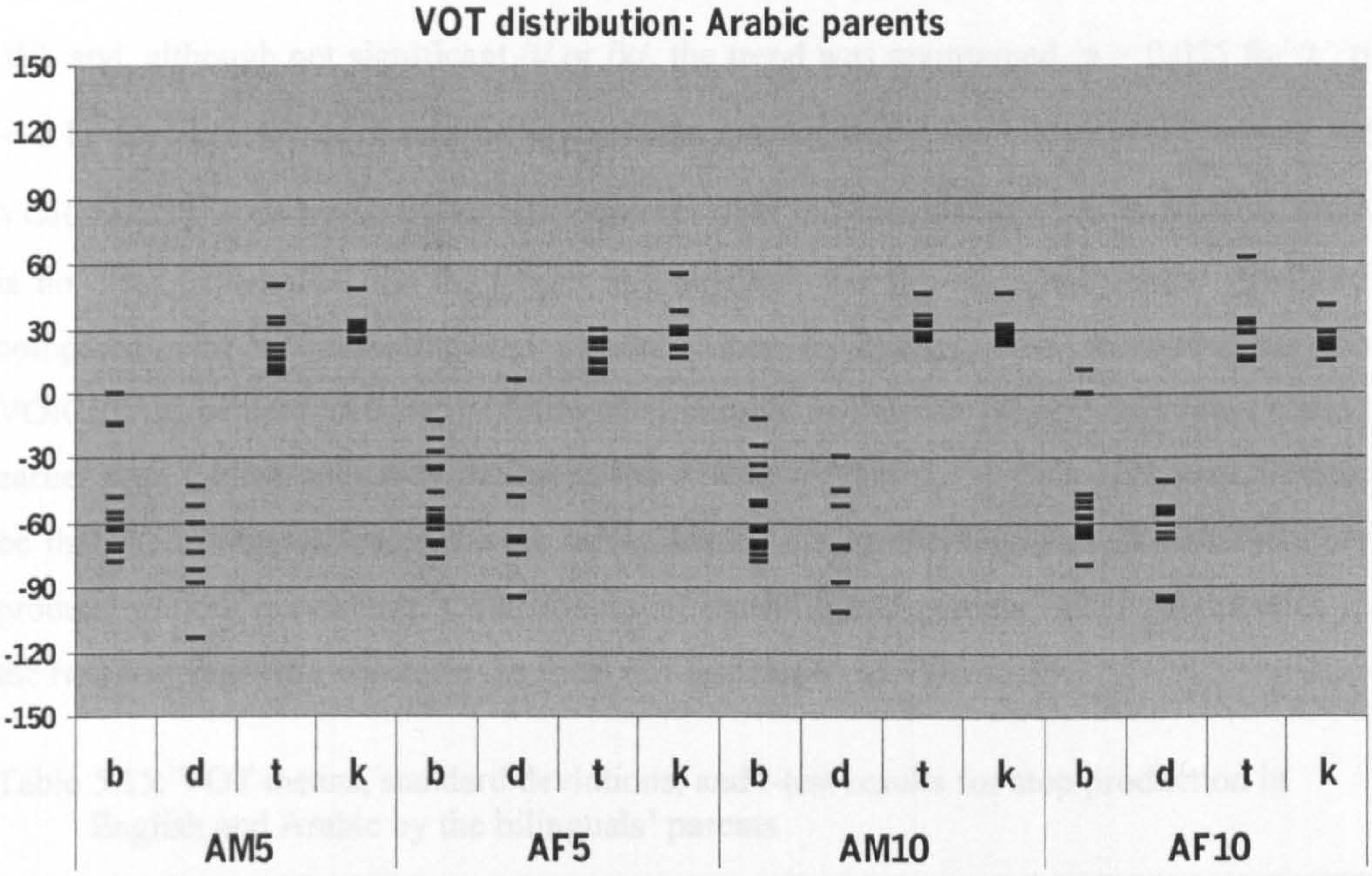


Figure 5.12: VOT distribution (in ms) for each of the Monolinguals' parents. N = 168.

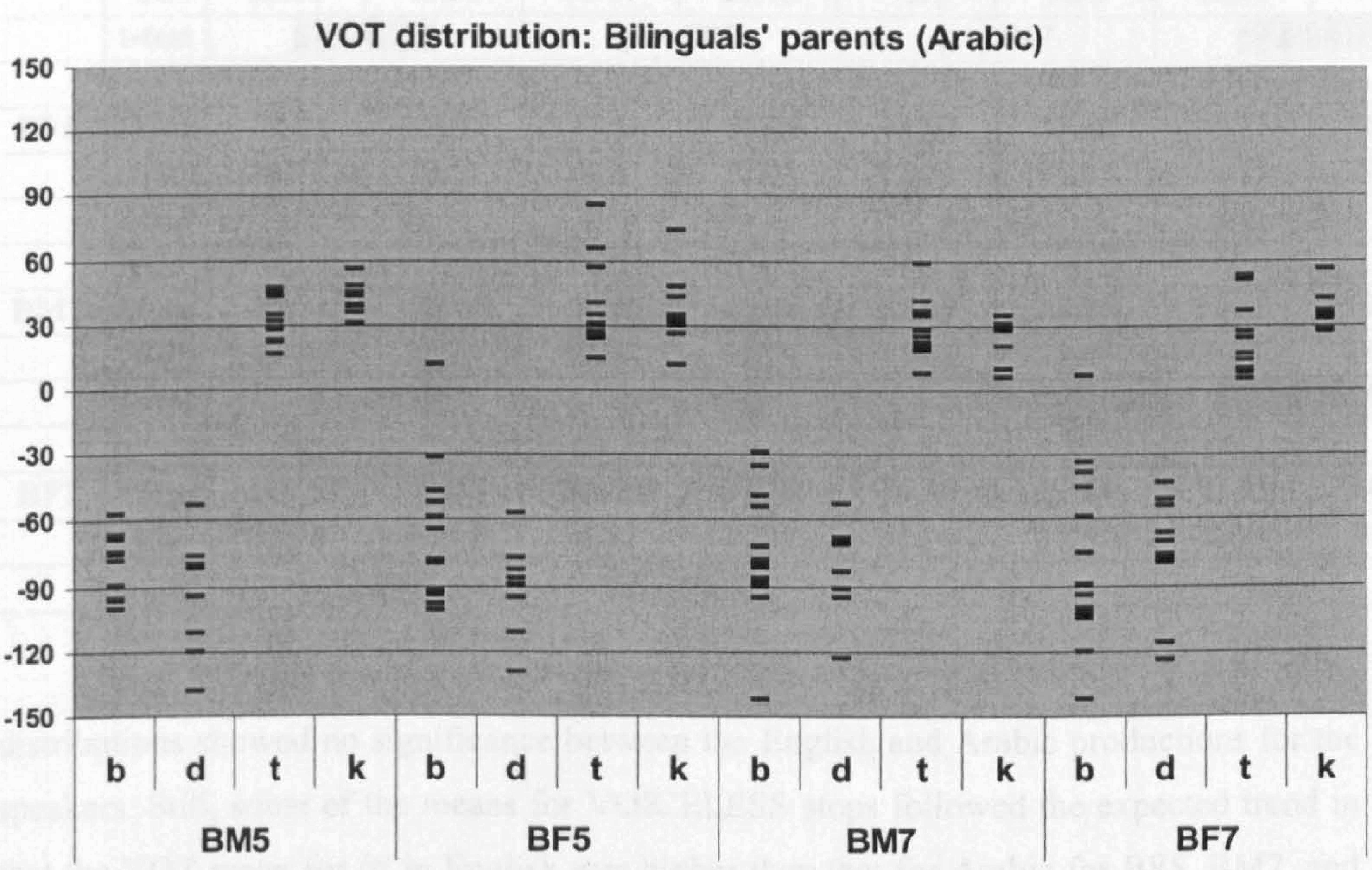


Figure 5.13: VOT distribution (in ms) for each of the bilinguals' parents in Arabic. N = 170

Overall, the patterns for the monolinguals' parents and the bilinguals' parents look similar, which is expected as Arabic is the L1 for the bilinguals' parents. However, there was a tendency for the bilinguals' parents to produce longer VOT values for both VOICED and VOICELESS stops than the monolinguals' parents. The difference was significant for the VOICED stops (t-test results significant at p << 0.001 for both /b/ and /d/), and, although not significant /t/ or /k/, the trend was maintained (p = 0.055 for /t/, p = 0.11 for /k/). While it can be argued that the tendency for higher VOT values for VOICELESS stops by the bilinguals' parents might indicate influence from English, there is no clear explanation for the longer voicing lead that the bilingual parents produced compared with the monolinguals' parents. More surprisingly, the prevoicing for the VOICED stops produced by the bilinguals' parents in English (Figure 5.6) often started earlier than for the ones they produced for Arabic (Figure 5.13). One explanation might be that the bilinguals' parents were trying harder during the English task and therefore producing more prevoicing. Comparisons of the bilingual parents' VOT productions in the two languages are shown in the form of t-test results in Table 5.15.

Table 5.15: VOT means, standard deviations, and t-test results for stop production in English and Arabic by the bilinguals' parents

		b		d		t		k	
		English	**Arabic**	**English**	**Arabic**	**English**	**Arabic**	**English**	**Arabic**
BM5	**Mean**	**-99.84**	**-78.67**	**-116.75**	**-94.38**	**30.60**	**30.83**	**33.40**	**42.22**
	SD	38.88	13.56	19.77	27.46	7.89	9.93	12.62	7.31
	t-test	p = .0206*		p = .085		p = .947		p= 0.041*	
BF5	**Mean**	**-74.91**	**-64.92**	**-77.56**	**-91.25**	**44.63**	**39.45**	**45.43**	**36.89**
	SD	49.74	30.51	18.30	22.08	13.50	21.35	16.79	17.24
	t-test	p = .536		p = .189		p = .488		p = .258	
BM7	**Mean**	**-78.74**	**-75.86**	**-92.88**	**-81.75**	**37.35**	**29.17**	**39.69**	**19.56**
	SD	42.08	28.13	23.72	21.59	15.63	13.48	9.08	11.73
	t-test	p = .805		p = .343		p = .144		p = 0.0007**	
BF7	**Mean**	**-118.52**	**-78.57**	**-105.56**	**-75.80**	**24.00**	**22.64**	**33.57**	**34.89**
	SD	32.98	40.11	31.81	26.37	11.96	17.57	8.38	9.41
	t-test	p = .005*		p = .0427*		p = .824		p = .737	

Apart from the few significant results highlighted in grey in Table 5.15, most of the distributions showed no significance between the English and Arabic productions for the speakers. Still, some of the means for VOICELESS stops followed the expected trend in that the VOT mean for /t/ in English was higher than that for Arabic for BF5, BM7, and BF7, and the VOT mean for /k/ in English was higher than that for Arabic for BF5 and significantly higher for BM7. However, the trend was sometimes reversed, in that the mean VOT for English /k/ by BM5 was actually significantly lower than that for Arabic /k/. On the whole, it seems that the bilinguals' parents are still applying their native VOT patterns onto their stop productions in both English and Arabic despite the extended exposure to English that they have all had since they have been living in the UK for more than ten years. The results for the bilinguals' parents in this study are different in this respect from comparable investigations by Flege (1987), Major (1992), and Sancier & Fowler (1997), who found significant changes in the production of adult bilinguals after several years of exposure to their L2. Reasons for the differences in results will be examined in Chapter Six.

Finally, in Section 5.9.1, it was noted that one of the bilinguals' parents (BF7) produced prevoiced stops in English that were often preceded by nasals and/or short central vowels (Figure 5.9). The same pattern was found for BF7's production in Arabic, e.g. [mbeːb] for [beːb] 'door', [nˈdawa] for [ˈdawa] 'medicine' and [ᵊmˈbɑʔˤɾa] for [ˈbɑʔˤɾa] 'cow' (Figure 5.14).

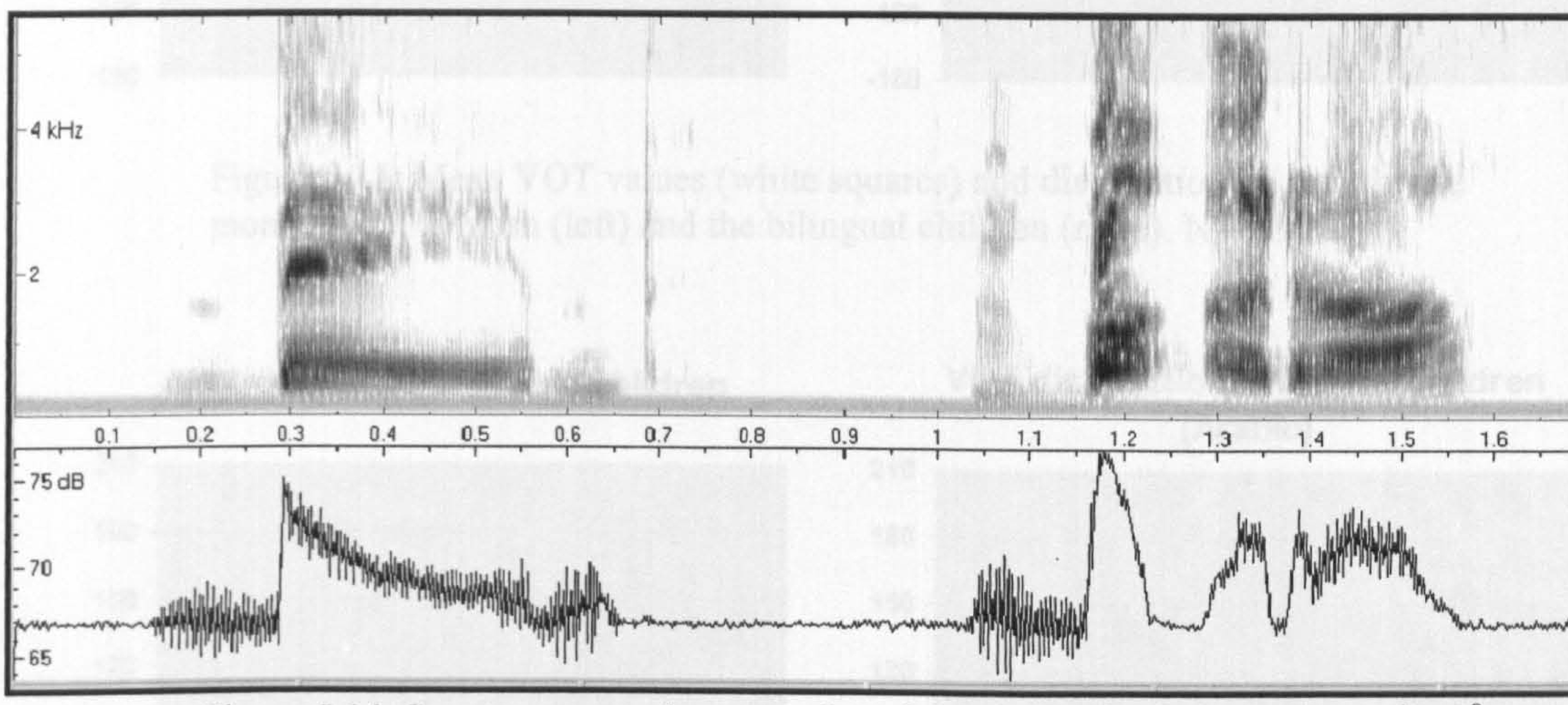


Figure 5.14: Spectrogram and energy plot of the words [beːb] 'door' and [ˈbɑʔˤɾa] 'cow' produced by BF7 as [mbeːb] and [ᵊmˈbɑʔˤɾa] respectively.

5.9.3 Children: group results

When the results for the three children in each language group are combined, the general pattern seems to suggest that the bilingual subjects are behaving differently in each language. However, while their English VOT patterns are similar to those of the monolingual English children, their Arabic patterns display some notable differences with those of the monolingual Arabic children. Figures 5.15 and 5.16 show mean VOT values and distributions for all the stops as produced by each group and in each language. /p/ and /g/ were included for Arabic because some of the children produced a number of loan words containing these sounds in initial position. Note that loan words are integrated into the phonology of Lebanese Arabic and are produced by all Lebanese speakers, not only the subjects in this study.

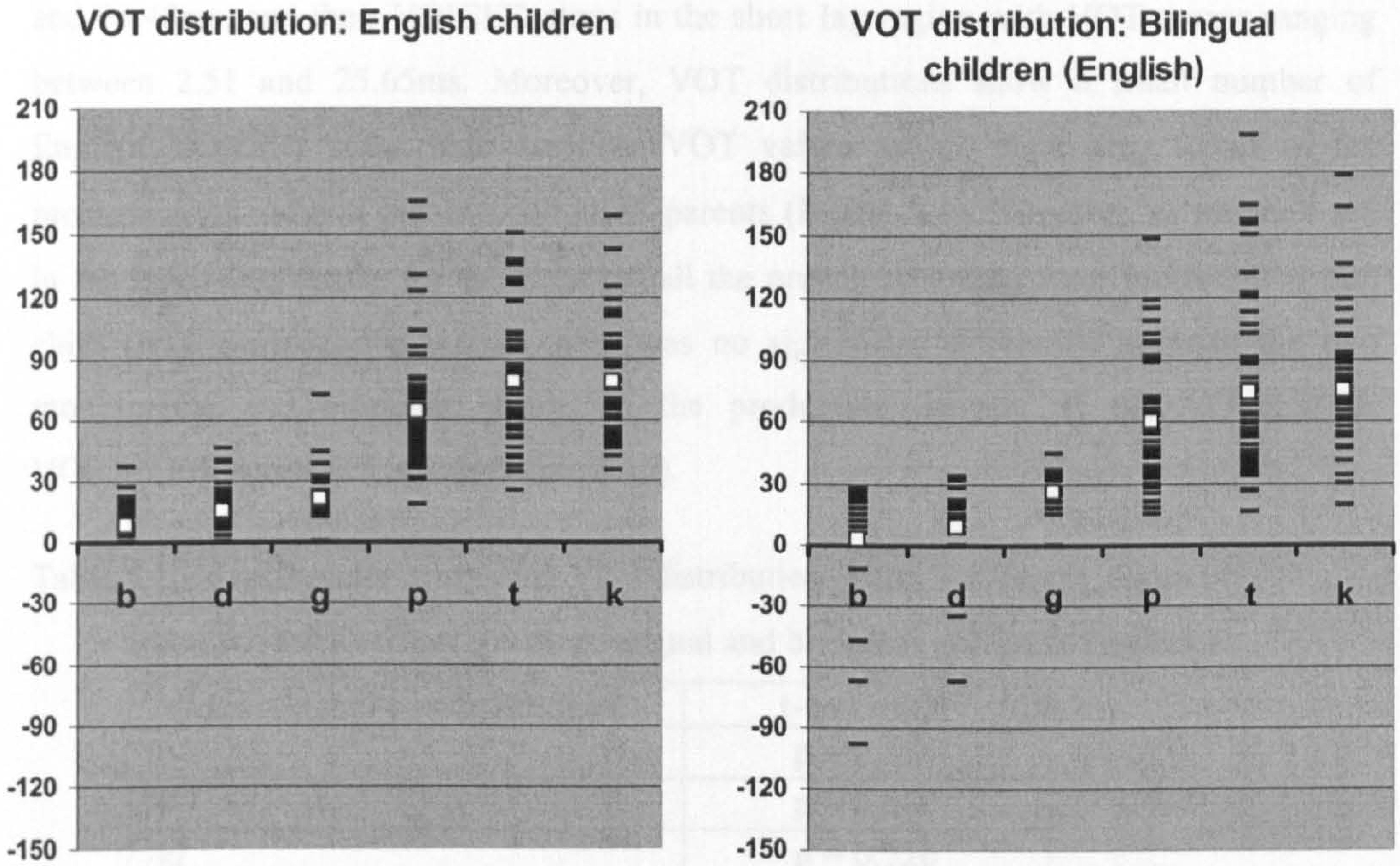


Figure 5.15: Mean VOT values (white squares) and distribution (in ms) for the monolingual English (left) and the bilingual children (right). N = 497.

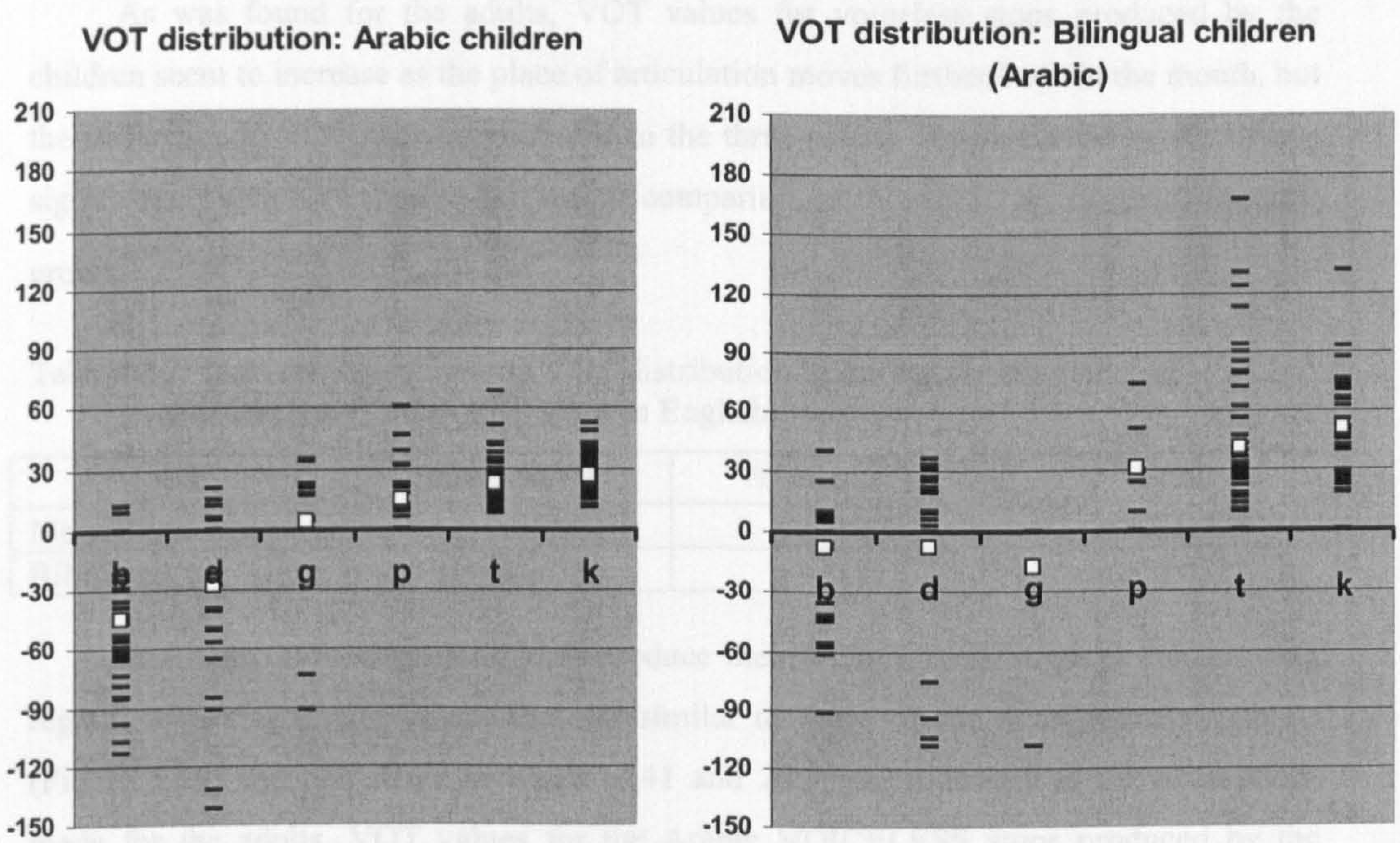


Figure 5.16: Mean VOT values (white squares) and distribution (in ms) for the monolingual Arabic (left) and the bilingual children (right). N = 365.

In English, both monolingual and bilingual subjects produce VOT values that are similar to adult values (Figure 5.15) and that are expected for this language. Their VOICELESS stops are in the long lag region with VOT means ranging between 59.94

and 78.48ms, and their VOICED stops in the short lag region with VOT means ranging between 2.51 and 25.65ms. Moreover, VOT distributions show a small number of English VOICED stops with negative VOT values which were also found in the production of some of the monolinguals' parents (Figure 5.5). However, as we shall see in the individual results for the children, all the prevoiced tokens were produced by one child (B7). Still, on the whole, there was no significant difference between the two monolingual and bilingual groups in the production of any of the VOICED or VOICELESS stops in English (Table 5.16).

Table 5.16: t-test results comparing VOT distribution in the 3 different places of articulation for each of the monolingual and bilingual groups in English.

Monolingual *versus* Bilingual	t-test result
/p/	p = 0.566
/t/	p = 0.540
/k/	p = 0.726
/b/	p = 0.080
/d/	p = 0.122
/g/	p = 0.195

As was found for the adults, VOT values for voiceless stops produced by the children seem to increase as the place of articulation moves further back in the mouth, but the difference in VOT distribution between the three places of articulation in not always significant. Table 5.17 shows t-test results comparing /p/, /t/ and /k/ distributions for each group.

Table 5.17: t-test results comparing VOT distribution in the 3 different places of articulation for VOICELESS stops in English.

	/p/ *versus* /t/	/t/ *versus* /k/	/p/ *versus* /k/
Monolinguals	p = 0.012*	p = 0.973	p = 0.021*
Bilinguals	p = 0.067	p = 0.813	p = 0.022*

In Arabic, monolingual subjects produce their VOICELESS stops in the short lag region, with mean VOT values that are similar to those of the monolinguals' parents (Figure 5.16) and that range between 14.41 and 26.93ms. Similarly to the observation made for the adults, VOT values for the Arabic VOICELESS stops produced by the monolingual Arabic children and those of the English VOICED stops produced by the monolingual English children seem to occupy slightly different ranges along the VOT continuum within the label 'short lag' although the difference is only significant for the alveolar place of articulation (Table 5.18).

Table 5.18: t-test results comparing VOT distribution between Arabic /p t k/ and English /b d g/ in the monolingual and bilingual children's productions.

Arabic monolinguals *versus* English monolinguals	t-test result
/p/ *versus* /b/	p = 0.127
/t/ *versus* /d/	p = 0.007*
/k/ *versus* /g/	p = 0.093
Bilingual Arabic /p/ *versus* Bilingual English /b/	p = 0.010*
Bilingual Arabic /t/ *versus* Bilingual English /d/	p = 0.0000***
Bilingual Arabic /k/ *versus* Bilingual English /g/	p = 0.0000***

Bilingual subjects had fewer prevoiced tokens for their VOICED stops than the monolingual Arabic children (20% for the bilinguals as opposed to 56% for the monolinguals), as well as significantly higher distributions for their Arabic VOICELESS stops than those of the monolinguals (results significant at $p \ll 0.01$ for /t/ and $p \ll 0.001$ for /k/[14]). Although this may suggest that there is influence from English on the bilingual's VOT production in Arabic, individual results discussed in 5.10.4 suggest that this does not apply to all three subjects. More importantly, the bilinguals' VOT distributions for Arabic VOICELESS stops are still markedly lower than those for English VOICELESS stops (results significant at $p \ll 0.001$ for both /t/ and /k/). Moreover, the bilinguals' Arabic VOICELESS stops and English VOICED stops (Figures 5.15 and 5.16) occupy slightly overlapping but markedly different ranges along the VOT continuum. However, while the VOT distributions of the monolingual English and Arabic VOICED stops were significantly different for /b/ and /d/, the bilinguals' distributions in English and Arabic struggled to reach significance (Table 5.19). Note that distributions for /g/ are insignificant between both monolingual groups and the bilinguals' production in the two languages due to the difficulty to maintain voicing lead in /g/ as explained in Section 5.9.1, which caused the /g/ productions by the children to fall in the short lag region regardless of the language in question. Moreover, for the monolingual children, /g/ occurs infrequently and in loan words only, which might be another reason why its VOT production is not mastered as well as that of /b/ and /d/.

Table 5.19: t-test results comparing VOT distribution between Arabic /b d g/ and English /b d g/ in the monolingual and bilingual children's productions.

	/b/	/d/	/g/
Monolingual English *versus* Arabic	p = 0.000**	p = 0.000**	p = 0.099
Bilingual English *versus* Arabic	p = 0.048*	p = 0.056	p = 0.060

[14] There were not enough /p/ tokens for each group to allow comparison.

While the VOT means for the Arabic VOICELESS stops increase as the place of articulation moves further back in the mouth for both monolingual and bilingual groups, the difference in the distribution between the three places of articulation is not always significant (Table 5.20).

Table 5.20: T-test results comparing VOT distribution in the 3 different places of articulation for VOICELESS stops in Arabic.

	/p/ *versus* /t/	/t/ *versus* /k/	/p/ *versus* /k/
Monolinguals	p = 0.053	p = 0.052	p = 0.003*
Bilinguals	p = 0.032*	p = 0.172	p = 0.057

So far we have been discussing the VOT results for the bilingual and monolingual children as two groups; we now move to the examination of individual results, as they show important findings that are related to age and the linguistic background of the children.

5.9.4 Children: individual results

In this section, each of the bilinguals' VOT production in the two languages is compared in order to examine whether the patterns adopted by each child and for each language are similar or different. The results are presented in increasing age order. The patterns for the monolingual subjects are also presented in each age group in order to allow for better interpretation of the results.

5.9.4.1 Age group 5:

In Figure 5.17, B5's VOT patterns in English and Arabic are presented against each other, followed by the patterns of each of the monolingual subjects of the same age (Figure 5.18). Means, ranges and standard deviations in English and Arabic for the five-year-old group are presented in Table 5.21.

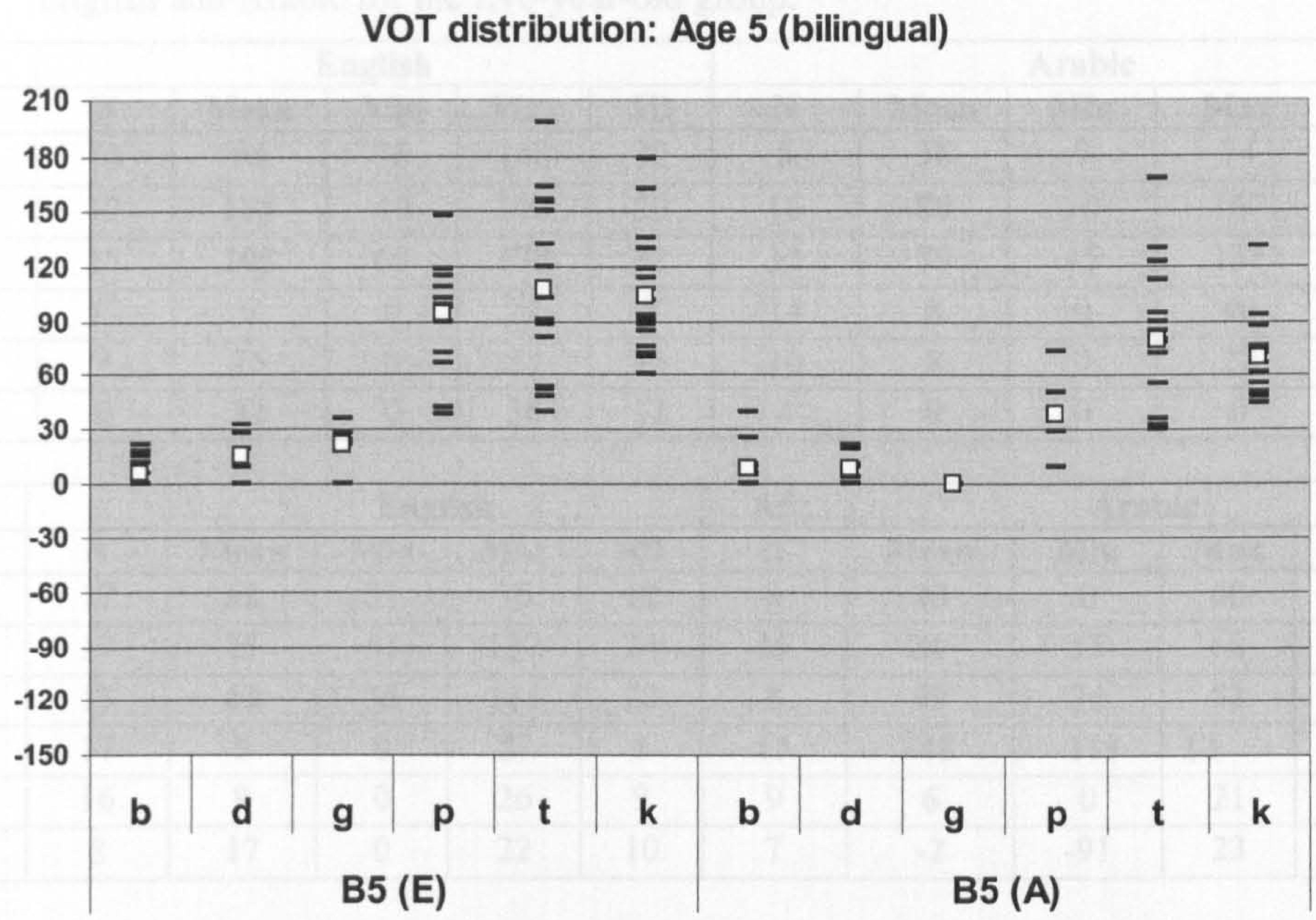


Figure 5.17: Mean VOT values (white squares) and distribution (in ms) for the B5 in English (left) and Arabic (right). N = 141.

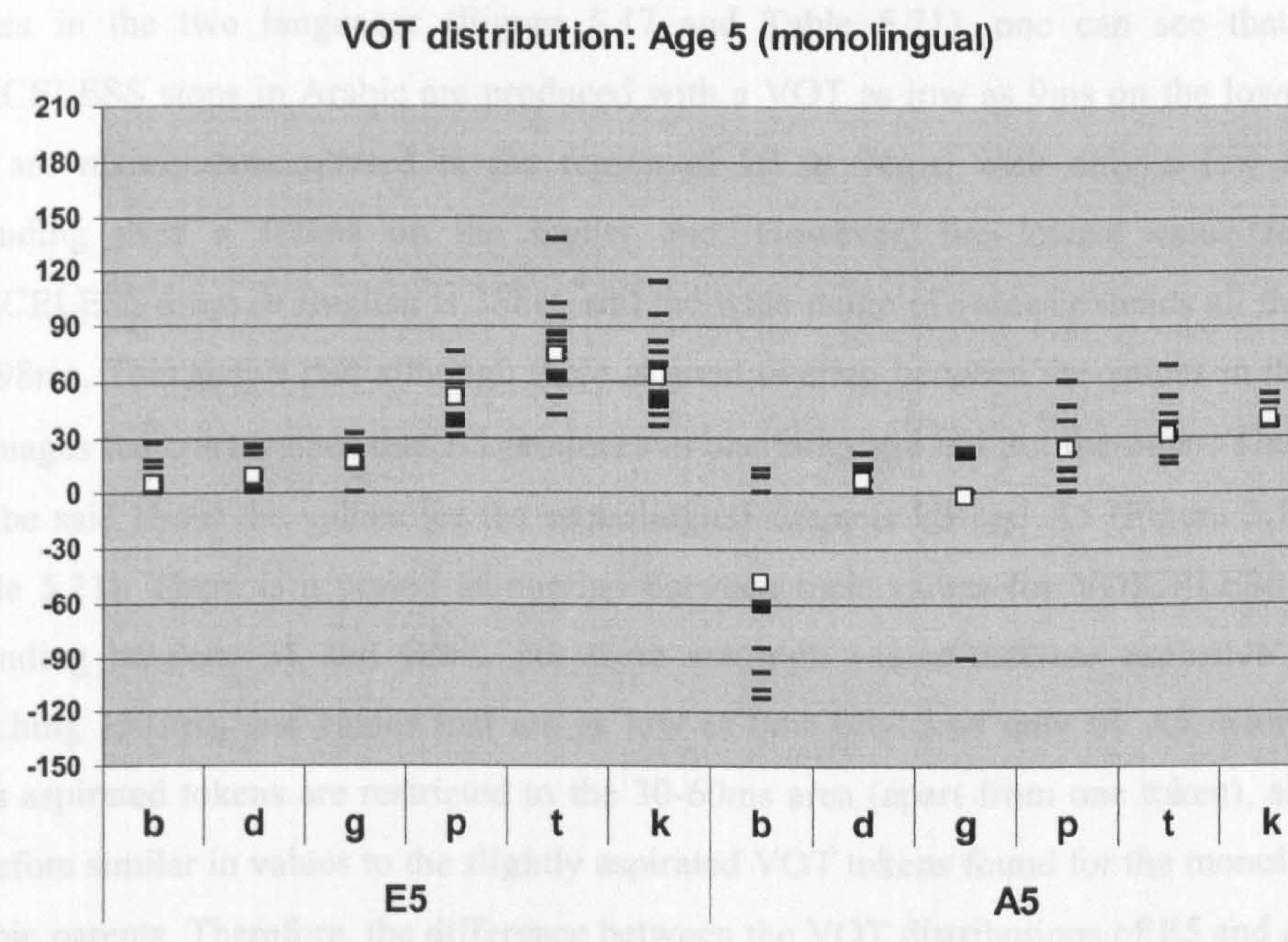


Figure 5.18: Mean VOT values (white squares) and distribution (in ms) for E5 (left) and A5 (right). N = 153.

Table 5.21: Detailed measurements of VOT means, ranges and standard deviations in English and Arabic for the five-year-old group.

B5	English					Arabic				
Stop	N	Mean	Min	Max	SD	N	Mean	Min	Max	SD
p	15	94	38	148	30	5	38	9	74	25
t	17	**108**	40	198	50	16	**80**	30	167	40
k	17	**105**	60	178	32	13	**70**	45	131	24
b	17	**6**	0	22	8	14	**8**	0	40	11
d	7	**15**	0	28	13	10	**8**	0	22	9
g	6	**22**	0	36	12	4	**0**	0	0	0

E5		English				A5	Arabic			
Stop	N	Mean	Min	Max	SD	N	Mean	Min	Max	SD
p	17	**52**	31	76	12	9	**23**	0	60	20
t	17	**75**	41	137	24	15	**30**	15	68	15
k	15	**62**	35	114	22	8	**40**	34	52	7
b	17	**5**	0	27	8	15	**-48**	-113	12	46
d	16	**9**	0	26	9	9	**6**	0	21	8
g	8	**17**	0	22	10	7	**-2**	-91	23	41

Looking at B5's results first, one can see that the VOICELESS stops in both languages are aspirated in her production. Still, t-tests show a statistical significance between the distributions of the VOT values in her production in the two languages for /p/ and /k/, though not for /t/ (Table 5.22). Moreover, looking at the ranges for the VOT values in the two languages (Figure 5.17 and Table 5.21), one can see that B5's VOICELESS stops in Arabic are produced with a VOT as low as 9ms on the lower end and are mainly concentrated in the region of 29 to 94ms, with only a few tokens extending over a 100ms on the higher end. However, the lowest value for her VOICELESS stops in English is 38ms, and the wide range of values extends all the way to 198ms. This shows that although there is great overlap between the values in the two languages there are values that B5 produces in one language but not the other. The same can be said about the values for the monolingual subjects E5 and A5 (Figure 5.18 and Table 5.21). There is a period of overlap between their values for VOICELESS stops extending between 31 and 68ms, but there are high values that are exclusive to E5 (reaching 137ms), and values that are as low as 0ms produced only by A5. Moreover, A5's aspirated tokens are restricted to the 30-60ms area (apart from one token), and are therefore similar in values to the slightly aspirated VOT tokens found for the monolingual Arabic parents. Therefore, the difference between the VOT distributions of E5 and A5 for VOICELESS stops is more significant than the difference between B5's distributions in English and Arabic (Table 5.22). Interestingly, B5's VOT values for VOICELESS stops in each language are in most cases significantly higher than the values produced by E5, the monolingual subject examined for that language.

Table 5.22: t-test result comparing the production of VOICELESS stops in English and Arabic by the bilingual and the monolingual 5-year-olds.

	/p/	/t/	/k/
B5 : English *versus* Arabic	p = 0.003*	p = 0.080	p = 0.003*
B5 (English) *versus* E5	p = 0.000**	p = 0.021	p = 0.000**
B5 (Arabic) *versus* A5	p = 0.272	p = 0.000**	p = 0.000**
E5 *versus* A5	p = 0.003*	p = 0.0000***	p = 0.001*

As for the VOICED stops, B5 uses short lag in both languages with considerable overlap between the values for /b/ and /d/. The only VOICED stop that is produced with different VOT ranges in each language is /g/, with a VOT of 0ms for all 4 tokens produced in Arabic while the English tokens ranged between 0 and 36ms. The distributions for /b/ and /d/ for B5 show no significant difference between English and Arabic (Table 5.23). However, it is important to note that the monolingual children also show some overlap in their production of VOICED stops, especially for /d/ (Figure 5.17). A5 still has not acquired voicing lead for /d/ and /g/, and produces them with short lag values that are similar to those produced by E5. However, A5 shows signs of acquisition of the voicing lead for /b/, with most of her values being produced in the long lead region (Table 5.21). The difference between the values for /b/ produced by the monolingual English and Arabic girls is significant (Table 5.23). Such results show a normal developmental pattern for A5, due to the fact that she is gradually acquiring voicing lead and starting with the place of articulation where it is easier to produce. B5, however, has not acquired voicing lead at any place of articulation.

Table 5.23: t-test result comparing the production of VOICED stops in English and Arabic by the bilingual and monolingual five-year-olds.

	/b/	/d/	/g/
B5 : English *versus* Arabic	p = 0.620	p = 0.246	p = 0.007*
B5 (English) *versus* E5	p = 0.670	p = 0.278	p = 0.489
B5 (Arabic) *versus* A5	p = 0.000**	p = 0.554	p = 0.922
E5 *versus* A5	p = 0.000**	p = 0.376	p = 0.269

In order to find out whether the VOT distributions for the Arabic VOICELESS stops are significantly different from those of the English VOICED stops, t-test results were run to compare A5's values for /p t k/ with E5's values for /b d g/ on the one hand, and B5's Arabic /p t k/ with her English /b d g/ on the other (Table 5.24).

Table 5.24: t-test results comparing VOT distribution between Arabic /p t k/ (A5) and English /b d g/ (E5) and between B5's Arabic /p t k/ and English /b d g/.

	/p/ versus /b/	*/t/ versus /d/*	*/k/ versus /g/*
A5 versus E5	p = 0.032*	p = 0.000**	p = 0. 000**
B5 (A) versus B5 (E)	p = 0.043*	p = 0. 0000***	p = 0. 000**

As can be seen for this age group the two sets of Arabic and English stops are significantly different from each other, with Arabic VOICELESS stops occupying higher ranges along the VOT continuum than English VOICED stops, despite the overlap (Figures 5.17 and 5.18).

In sum, B5's VOT patterns for VOICELESS stops in English and Arabic are on the whole significantly different from each other, but her patterns for VOICED stops in the two languages show considerable overlap and are not significantly different. In English, her VOT production follows the expected pattern of short lag/long lag, often with particularly high values for her VOICELESS stops. In Arabic, however, her production does not follow the expected long lead/short lag pattern. Her VOICELESS stops are mainly aspirated, while her VOICED stops are produced with a short lag. However, there are two important things to consider before attributing those results to the fact that B5 is bilingual. With regards to her use of short lag for her VOICED stops in Arabic, note that A5 shows similar patterns and has not acquired voicing lead for all her stops. This is most likely due to the subjects' young age and the difficulties related to the production of voicing lead that were discussed in Section 5.5.2. As for B5's production of long lag VOT for her VOICELESS stops in Arabic, note that her VOT values for VOICELESS stops are generally high even in English, and this may be due to her young age, but also to the fact that she has had less exposure to Arabic than A5. What is important though, is that B5 still produces significantly higher VOT values for two of her VOICELESS English stops than the Arabic counterparts.

5.9.4.2 Age group 7:

In Figure 5.19 B7's VOT patterns in English and Arabic are presented against each other, followed by the patterns of the two monolingual subjects of the same age (Figure 5.20). Means, ranges and standard deviations in English and Arabic for the seven-year-old group are presented in Table 5.25.

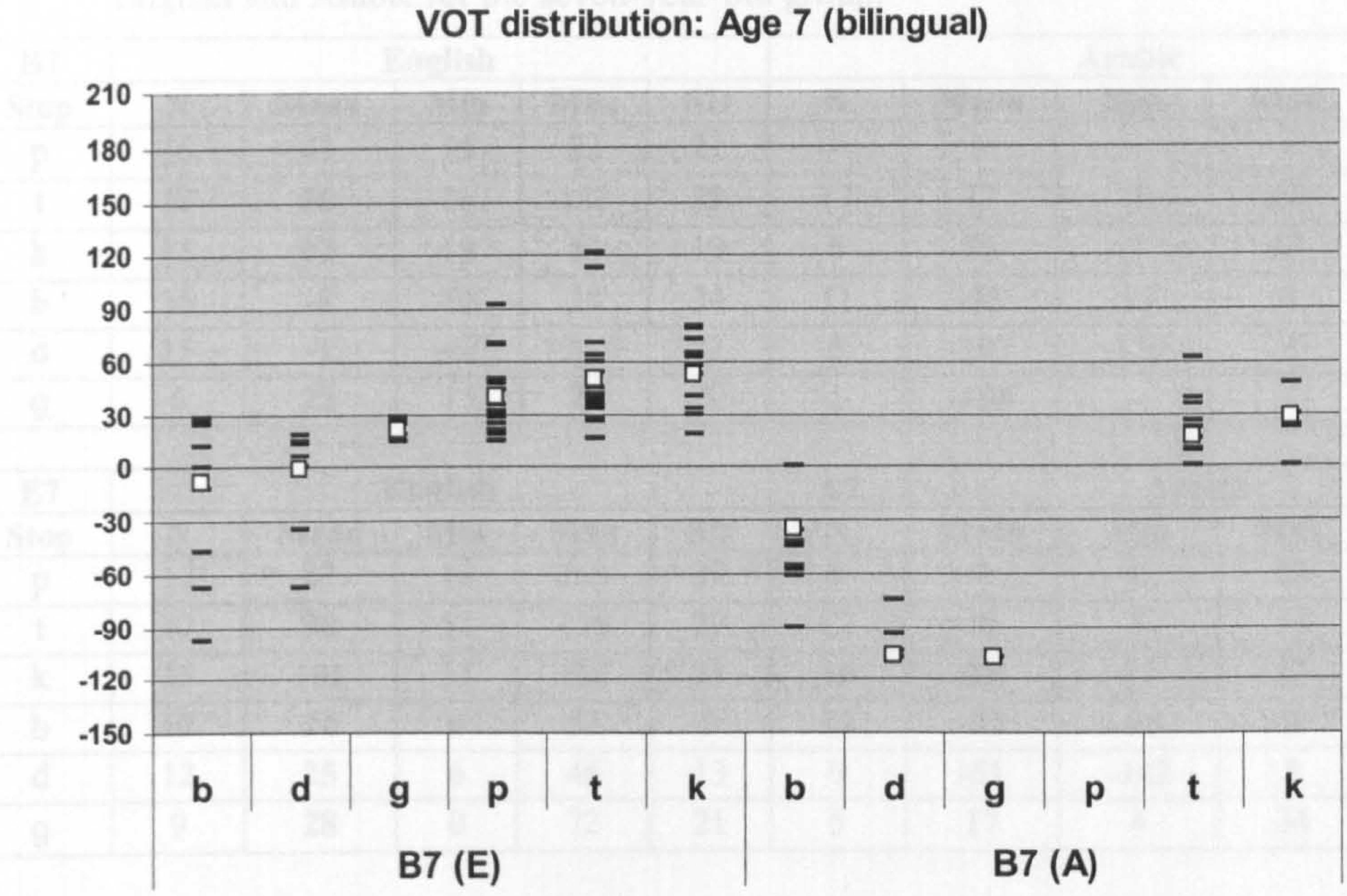


Figure 5.19: Mean VOT values (white squares) and distribution (in ms) for the B7 in English (left) and Arabic (right). N = 125.

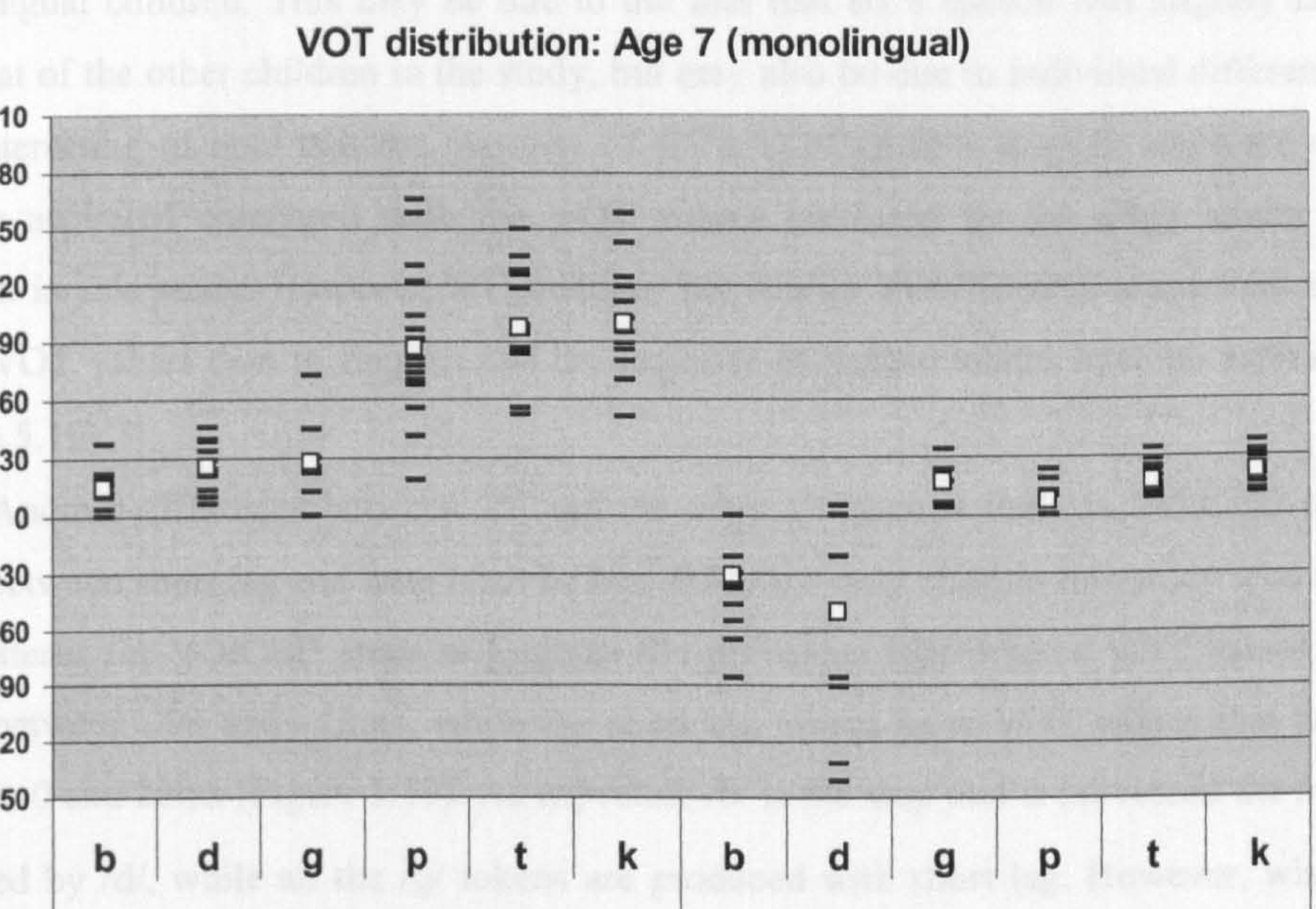


Figure 5.20: Mean VOT values (white squares) and distribution (in ms) for E7 (left) and A7 (right). N = 157.

Table 5.25: Detailed measurements of VOT means, ranges, and standard deviations in English and Arabic for the seven-year-old group.

B7	English					Arabic				
Stop	N	Mean	Min	Max	SD	N	Mean	Min	Max	SD
p	16	**41**	15	92	21	-	-	-	-	-
t	17	**50**	16	122	29	17	**17**	0	62	19
k	15	**52**	19	80	19	6	**28**	0	48	18
b	16	**-8**	-98	28	34	11	**-35**	-91	0	31
d	15	**-1**	-67	19	23	5	**-107**	-152	-75	28
g	6	**22**	15	28	5	1	**-108**	-	-	-

E7	English					A7	Arabic			
Stop	N	Mean	Min	Max	SD	N	Mean	Min	Max	SD
p	17	**87**	18	165	39	9	**7**	0	23	9
t	17	**98**	52	150	29	17	**16**	8	34	7
k	15	**101**	51	158	31	16	**21**	11	37	8
b	17	**14**	0	37	9	14	**-32**	-86	0	30
d	12	**25**	6	46	13	9	**-51**	-142	5	62
g	9	**28**	0	72	21	5	**17**	4	34	12

Looking at B7's results first, one can see that, unlike B5, there is a significant difference between the VOT values for both his VOICELESS and VOICED stops in English and Arabic (Table 5.26). In English, B7's VOICELESS stops are aspirated, though not as highly as any of the stops produced for English by the other bilingual and monolingual children. This may be due to the fact that B7's speech was slightly faster than that of the other children in the study, but may also be due to individual differences. It is interesting to note that the majority of B7's VOICELESS English stops are only slightly aspirated compared with the VOT values produced by the other adults and children in this study. However, B7 produces his Arabic VOICELESS stops with even lower VOT values than in English and the majority of Arabic tokens have no aspiration (Figure 5.19).

Another difference between B7 and the other children is that his VOICED stops range between short lag and long lead. In fact, B7 is the only child in this study who uses two patterns for VOICED stops in English: the prevoiced tokens have VOT values that range between –98 and –12ms, while the short lag tokens have VOT values that range between 0 and 28ms (Figure 5.19). As expected, /b/ is the stop that is prevoiced the most, followed by /d/, while all the /g/ tokens are produced with short lag. However, what is particular about B7's prevoiced stops in English is that in six out of the nine prevoiced tokens, the voicing that precedes the release of the stop is accompanied by energy above the F0 level (Figure 5.21) that is similar to the type found for his mother (BF7) in Section 5.9.1, though only one token had audible nasality (but cf. Arabic results in the next paragraph). Instead, on closer auditory inspection, three tokens were actually produced as implosives [ɓ] (e.g. [ˈɓʌt^{h}ə] for 'butter').

In Arabic, no /p/ tokens were available for analysis, but the other two VOICELESS stops /t/ and /k/ were produced with short lag to slight aspiration, with values ranging between 0 and 68ms (Table 5.25). As for the Arabic VOICED stops, they are mainly produced with voicing lead, though a few /b/ tokens have a value of 0ms. Most of the tokens are produced in the region of -160 and -40ms, with a few tokens between -40 and 0ms (Figure 5.19). Moreover, some of B7's Arabic prevoiced stops, like his English prevoiced stops, have features that were noted in the analysis of his mother's speech (Section 5.9.1). Auditory analysis reveals that 11 out of the 24 /b/ and /d/ tokens in Arabic sound like [mb] and [nd] respectively, and spectrographic analysis reveals that the laryngeal voicing preceding these stops is accompanied by traces of formant structure (Figure 5.22). The only other child who produces similar features is the five-year-old Sarah, but for only two out of the ten prevoiced tokens.

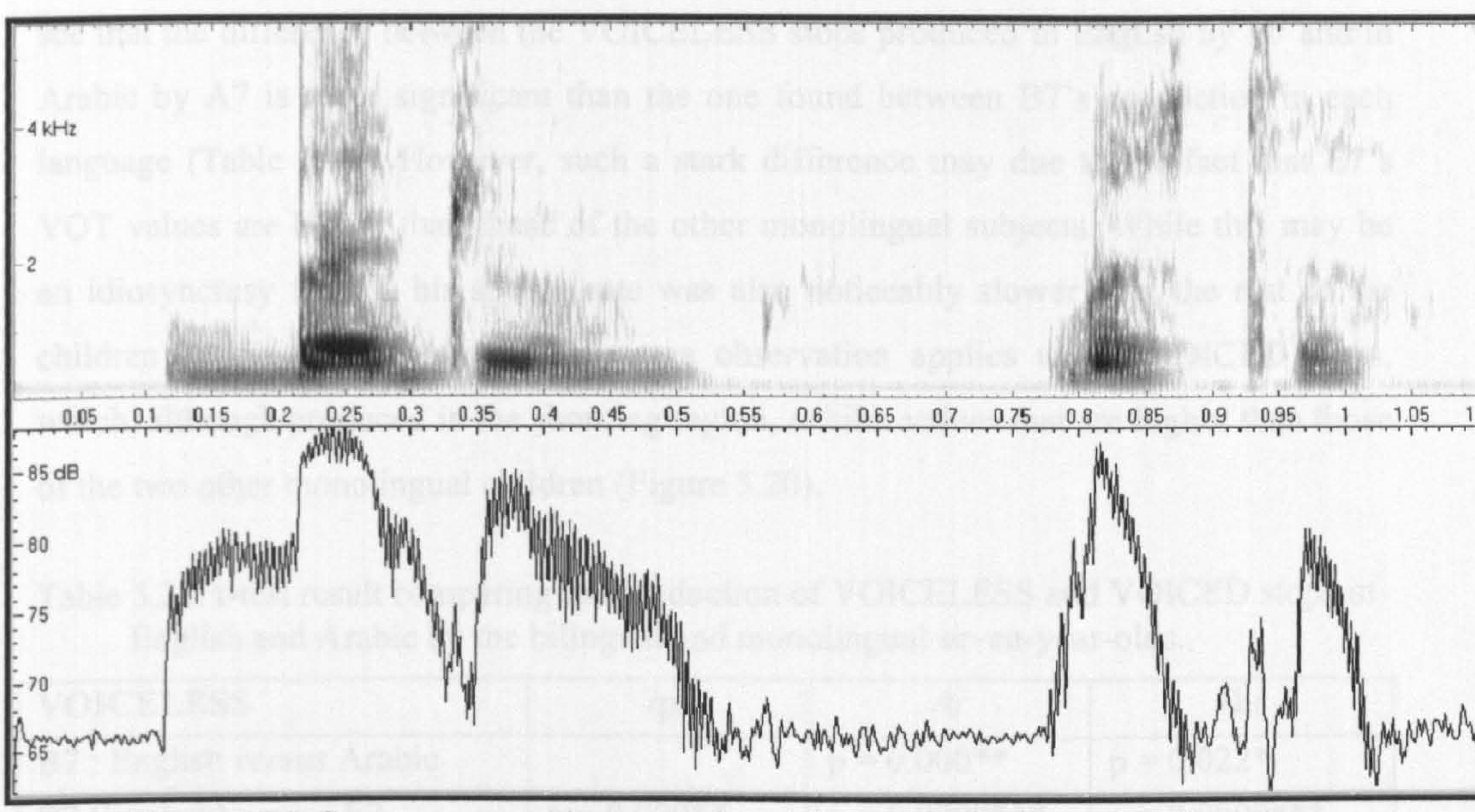


Figure 5.21: Spectrogram and energy plot of the words 'bedroom' (left) and 'butter' (right) as produced by B7 as [^{ə}mˈbɛd^{ʒ}ɹʉm] and [ˈɓʌt^{h}ə] respectively.

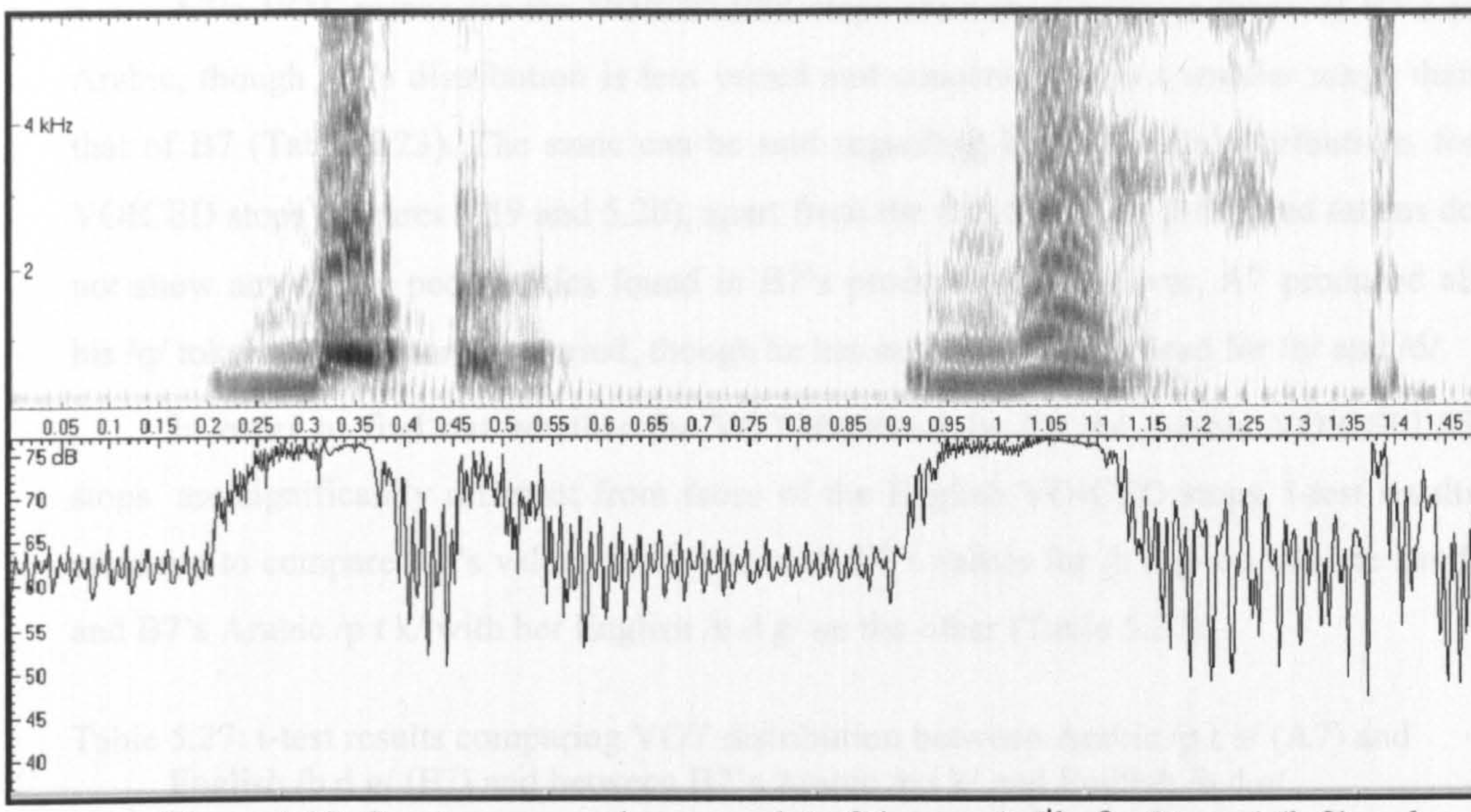


Figure 5.22: Spectrogram and energy plot of the words [ˈbɑʔɾɐ] 'cow' (left) and [di:k] 'cockerel' (right) as produced by B7, showing formant structure and energy above the F0 level.

Looking at the patterns of the monolingual seven-year-olds (Figure 20), one can see that the difference between the VOICELESS stops produced in English by E7 and in Arabic by A7 is more significant than the one found between B7's production in each language (Table 5.26). However, such a stark difference may due to the fact that E7's VOT values are higher than those of the other monolingual subjects. While this may be an idiosyncrasy for E7, his speech rate was also noticeably slower than the rest of the children during the recordings. The same observation applies to his VOICED stops, which, although produced in the short lag region, exhibit values that are higher than those of the two other monolingual children (Figure 5.20).

Table 5.26: t-test result comparing the production of VOICELESS and VOICED stops in English and Arabic by the bilingual and monolingual seven-year-olds.

VOICELESS	/p/	/t/	/k/
B7 : English *versus* Arabic		p = 0.000**	p = 0.022*
B7 (English) *versus* E7	p = 0.000**	p = 0.0000***	p = 0.0000***
B7 (Arabic) *versus* A7		p = 0.896	p = 0.386
E7 *versus* A7	p = 0.0000***	p = 0.0000***	p = 0.0000***
VOICED	/b/	/d/	/g/
B7 : English *versus* Arabic	p = 0.041*	p = 0.000**	
B7 (English) *versus* E7	p = 0.026*	p = 0.001*	p = 0.419
B7 (Arabic) *versus* A7	p = 0.787	p = 0.040*	
E7 *versus* A7	p = 0.0000***	p = 0.006*	p = 0.246

A7's VOT means for the VOICELESS stops are very similar to those of B7's in Arabic, though A7's distribution is less varied and concentrated in a smaller range than that of B7 (Table 5.23). The same can be said regarding B7 and A7's distributions for VOICED stops (Figures 5.19 and 5.20), apart from the fact that A7's prevoiced tokens do not show any of the peculiarities found in B7's production. Moreover, A7 produced all his /g/ tokens in the short lag period, though he has acquired voicing lead for /b/ and /d/.

In order to find out whether the VOT distributions for the Arabic VOICELESS stops are significantly different from those of the English VOICED stops, t-test results were run to compare A7's values for /p t k/ with E7's values for /b d g/ on the one hand, and B7's Arabic /p t k/ with her English /b d g/ on the other (Table 5.27).

Table 5.27: t-test results comparing VOT distribution between Arabic /p t k/ (A7) and English /b d g/ (E7) and between B7's Arabic /p t k/ and English /b d g/.

	/p/ *versus* /b/	/t/ *versus* /d/	/k/ *versus* /g/
A7 versus E7	p = 0.068	p = 0.039*	p = 0.385
B7 (A) versus B7 (E)		p = 0.026*	p = 0.416

As opposed to the findings for the five-year-old group (Table 5.24) where the three places of articulation were significantly different for both language groups, only the distributions for Arabic /t/ and English /d/ turned out to be significantly different for the seven-year-old group. The reason behind this is that both A7 and B7 produce VOT values for Arabic VOICELESS stops that are lower than expected, as they tend to be concentrated in the 0-30ms zone of the short lag region rather than spanning to the 30-60ms region as was found for the adults (Figure 5.10).

In sum, B7's VOT patterns in English and Arabic are significantly different for both VOICED and VOICELESS stops. In English, he produces long lag for VOICELESS stops and both long lead and short lag for his VOICED stops. In Arabic, he produces short lag for his VOICELESS stops, and mainly long lead for his VOICED stops. B7's Arabic VOT patterns are closer to those of the monolingual child his age (A7) than his English patterns are to those of E7. On the one hand, this is evident in the fact that both B7 and A7 produce short VOT values for their Arabic VOICELESS stops and that B7 shows signs of having acquired voicing lead. On the other hand, B7 produces shorter VOT values for his VOICELESS stops in English than any of the children in the study, while E7 produces particularly long VOT values, which increases the difference between B7 and E7. Finally, B7's prevoiced tokens have peculiar auditory and spectrographic features that are not found in the production of the other children but that are found in his mother's production.

5.9.4.3 Age group 10:

In Figure 5.23, B10's VOT patterns in English and Arabic are presented against each other, followed by the patterns of the two monolingual subjects of the same age (Figure 5.24). Means, ranges and standard deviations in English and Arabic for the five-year-old group are presented in Table 5.28.

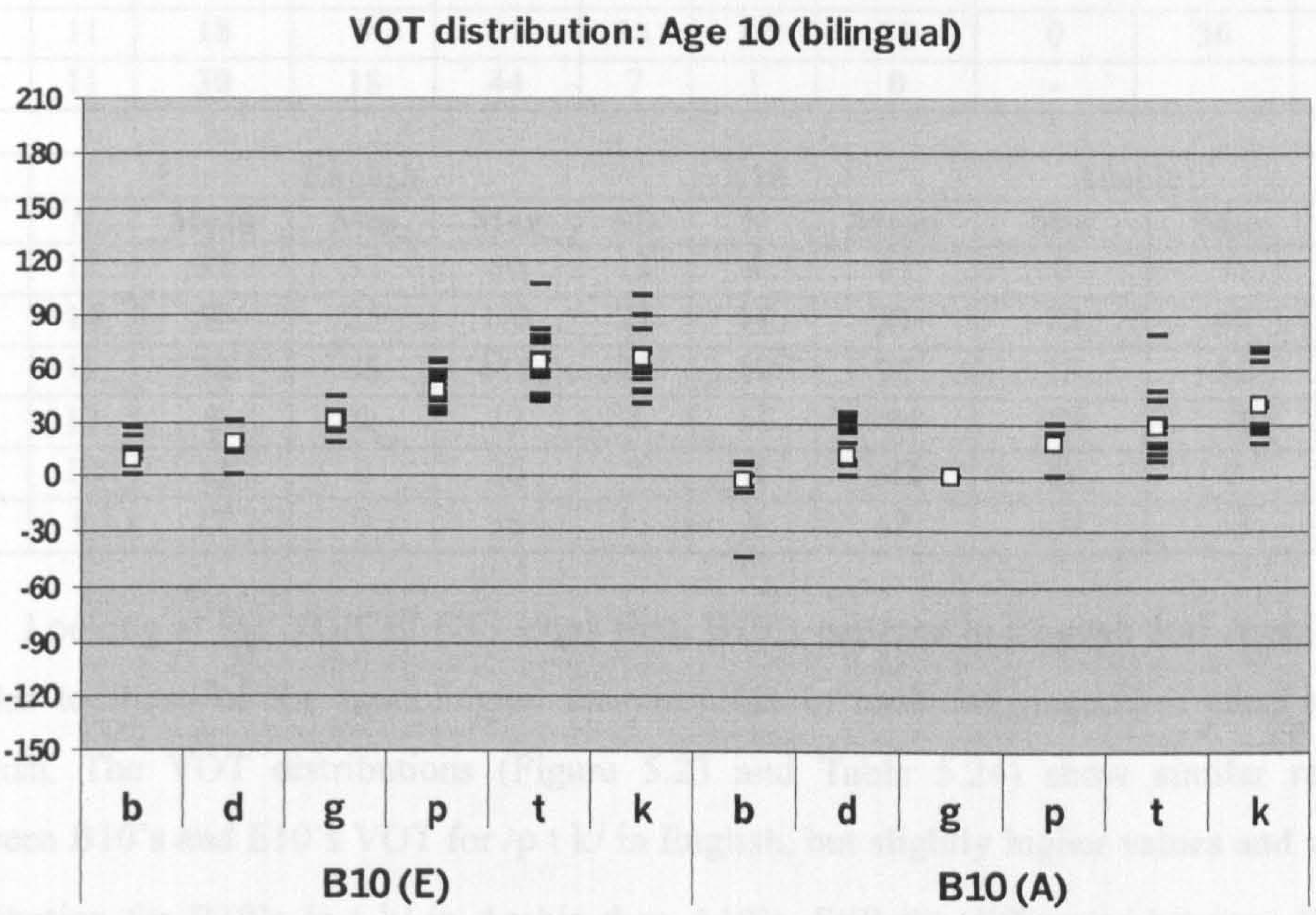


Figure 5.23: Mean VOT values (white squares) and distribution (in ms) for the B10 in English (left) and Arabic (right). N = 149.

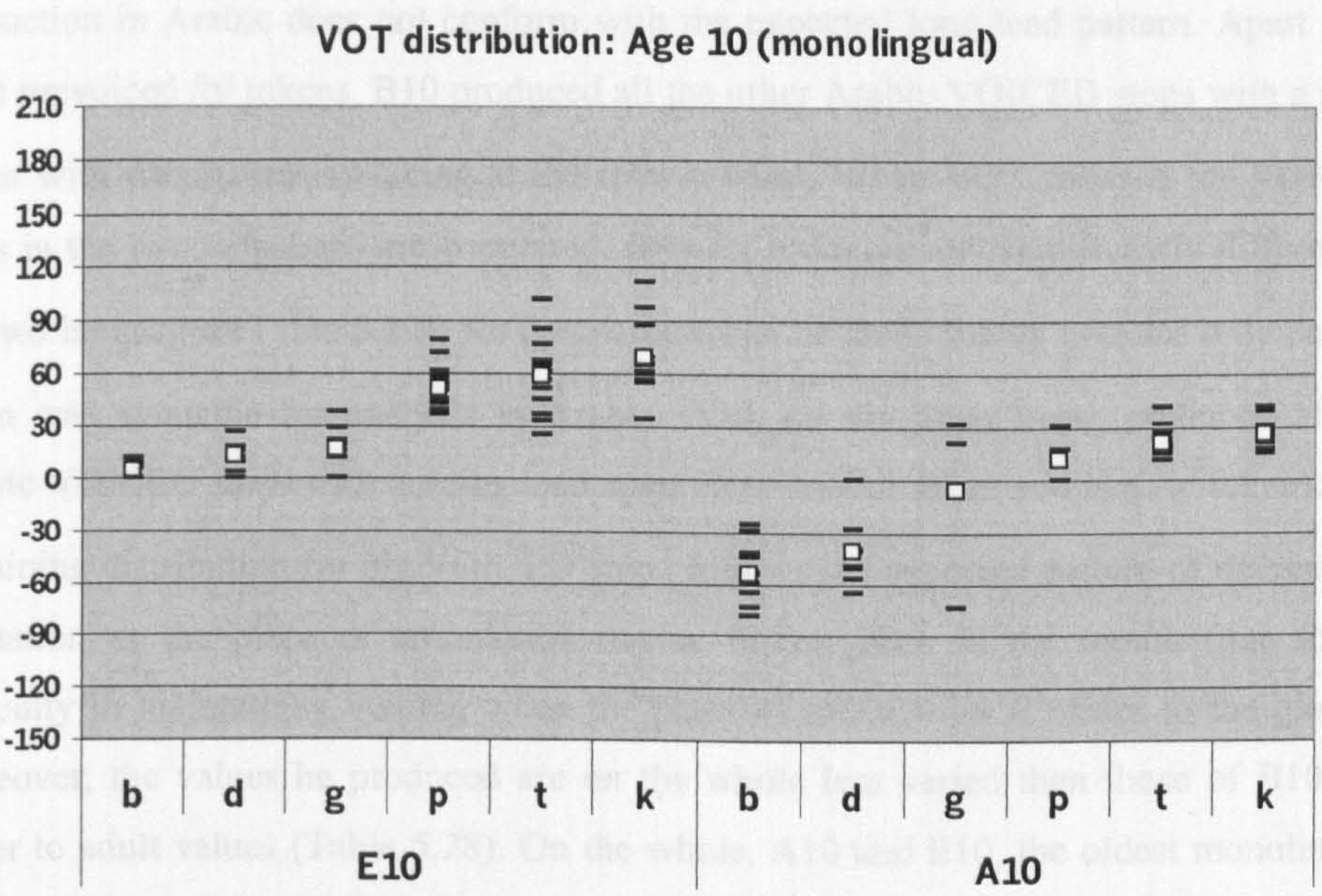


Figure 5.24: Mean VOT values (white squares) and distribution (in ms) for E10 (left) and A10 (right). N = 137.

Table 5.28: Detailed measurements of VOT means, ranges, and standard deviations in English and Arabic for the ten-year-old group

B10	English					Arabic				
Stop	N	Mean	Min	Max	SD	N	Mean	Min	Max	SD
p	17	48	34	65	10	3	18	0	30	16
t	16	63	42	107	17	17	28	0	79	19
k	13	66	40	101	20	11	41	18	72	20
b	16	9	0	27	8	16	-2	-44	8	12
d	11	18	0	31	11	17	11	0	36	14
g	11	30	18	44	7	1	0	-		-

E10	English					A10	Arabic			
Stop	N	Mean	Min	Max	SD	N	Mean	Min	Max	SD
p	17	52	37	80	12	4	12	0	31	15
t	14	59	25	103	21	17	21	12	42	8
k	11	70	35	112	22	17	27	15	40	8
b	12	5	0	12	4	16	-54	-25	-79	16
d	10	13	0	26	7	8	-42	-66	0	20
g	8	17	0	38	13	3	-7	-74	33	59

Looking at the VOICELESS stops first, B10's patterns in English and Arabic are similar to those of the monolingual ten-year-olds in each language, but more so in English. The VOT distributions (Figure 5.23 and Table 5.24) show similar ranges between B10's and E10's VOT for /p t k/ in English, but slightly higher values and wider distribution for B10's /p t k/ in Arabic than A10's. Still the difference between B10's VOT for VOICELESS stops in the two languages is significant (Table 5.29). As for the VOICED stops, B10's VOT patterns in English are similar to those of E10, but his VOT production in Arabic does not conform with the expected long lead pattern. Apart from three prevoiced /b/ tokens, B10 produced all the other Arabic VOICED stops with a short lag or with simultaneous voicing at the release burst. When VOT patterns for VOICED stops in the two languages are compared, only /b/ production is significantly different in the two languages (Table 5.29). No comparison can be made for /g/ because only one /g/ token was available for analysis in Arabic. A10, on the other hand, produced all his Arabic VOICED stops with voicing lead apart from one /d/ token and two /g/ tokens. The remaining distribution for his VOICED stops follows the expected pattern of decrease in phonation as the place of articulation moves further back in the mouth (due to the difficulty in maintaining voicing when the place of articulation is closer to the glottis). Moreover, the values he produced are on the whole less varied than those of B10 and closer to adult values (Table 5.28). On the whole, A10 and E10, the oldest monolingual children in this study, are the ones who produced VOT patterns that are more stable and more similar to monolingual adult patterns than the rest of the bilingual and monolingual children.

Table 5.29: t-test result comparing the production of VOICELESS and VOICED stops in English and Arabic by the bilingual and monolingual ten-year-olds.

VOICELESS	/p/	/t/	/k/
B10 : English *versus* Arabic		p = 0. 0000***	p = 0.006*
B10 (English) *versus* E10	p = 0.238	p = 0.641	p = 0.636
B10 (Arabic) *versus* A10	p = 0.652	p = 0.208	p = 0.047*
E10 *versus* A10	p = 0.009*	p = 0.0000***	p = 0.0000***
VOICED	/b/	/d/	/g/
B10 : English *versus* Arabic	p = 0.006*	p = 0.141	
B10 (English) *versus* E10	p = 0.158	p = 0.181	p = 0.027*
B10 (Arabic) *versus* A10	p = 0.0000***	p = 0.0000***	
E10 *versus* A10	p = 0.001*	p = 0.025*	

In order to find out whether the VOT distributions for the Arabic VOICELESS stops are significantly different from those of the English VOICED stops, t-test results were run to compare A10's values for /p t k/ with E10's values for /b d g/ on the one hand, and B10's Arabic /p t k/ with her English /b d g/ on the other (Table 5.30).

Table 5.30: t-test results comparing VOT distribution between Arabic /p t k/ (A10) and English /b d g/ (E10), and between B10's Arabic /p t k/ and English /b d g/.

	/p/ *versus* /b/	/t/ *versus* /d/	/k/ *versus* /g/
A10 versus E10	p = 0.420	p = 0.008*	p = 0.087
B10 (A) versus B5 (E)		p = 0.101	p = 0.087

Although there is a definite tendency for the Arabic means to be higher than the English ones for both B10 (Arabic) *versus* B10 (English) and A10 *versus* E10 (Table 5.28), the VOT distributions for the two sets of stops show no significant difference, apart from the difference between Arabic /t/ and English /d/ in the production of the monolingual children. Since similar results were found for the seven-year-olds (Table 5.27), this confirms that there are other acoustic and articulatory cues that the children are using to yield the auditory quality of /p t k/ on the one hand for Arabic, and /b d g/ on the other for English.

In sum, B10 produces VOT values that are more stable than those produced by the two younger bilingual subjects in this study. His VOT patterns show closer resemblance to the monolingual children of his age and to the expected adult patterns in each language. However, the only pattern he has not acquired yet is voicing lead for his

VOICED stops in Arabic, and it is surprising that his younger brother, B7, is the only bilingual subject who shows signs of acquisition of the voicing lead. Possible reasons for the difference in the speech behaviour of the two brothers will be presented in Section 5.9.6.

5.9.5 Summary of children's results

Figures 5.25 and 5.26 show the mean VOT values for each of the English and Arabic monolingual and bilingual subjects. Starting with the monolingual children, in English, the three subjects produce their stops according to the expected short lag/long lag pattern. Detailed results (Section 5.9.4) show a gradual progression towards adult models in that the VOT values for the stops become shorter and less variable with age. E10 has the smallest standard deviations for the means of all his stops (5.9.4.3), whereas E5 has the highest standard deviations and ranges of VOT values (5.9.4.1). Individual differences that are not necessarily related to age include the fact that E7 has higher VOT values for both his VOICED and VOICELESS stops than the other two subjects, which may partly be the result of the slower rate of his speech compared with the rest of the children.

In Arabic, there are also signs of gradual progression towards adult values for the monolingual subjects in that A10 is the only subject who has acquired full voicing lead for all his VOICED stops and short lag for his VOICELESS stops. Moreover, his VOT values are on the whole less variable than those of the two younger subjects (Section 5.9.4). A7 and A5 have wider VOT ranges and show incomplete acquisition of the voicing lead, as some of their VOICED stops are produced with a short lag. Note, however, that /g/ is the only stop for which A7 still has not acquired a voicing lead production, while A5 still needs to acquire voicing lead for /g/ and /d/. Finally, although the values obtained for the VOICELESS stops by the three monolingual subjects are generally higher than those obtained by the monolingual English subjects for their VOICED stops, the difference is not always significant due to the considerable overlap between the two sets of stops. The values for the VOICELESS stops produced by the Arabic monolingual children are generally lower than the ones found for their parents, and suggest the children may be using other cues for achieving the perceptual quality required for Arabic /p t k/ and English /b d g/.

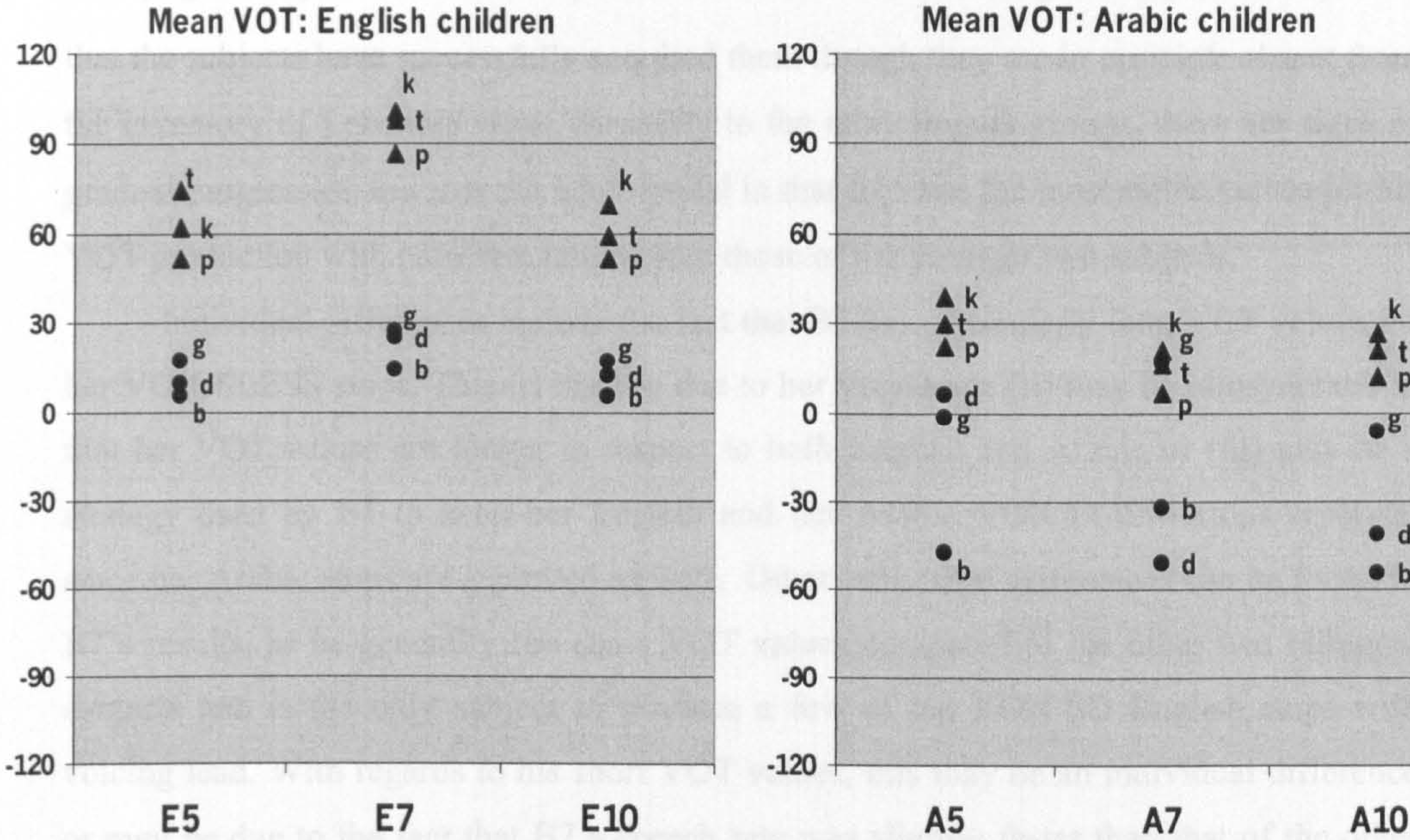


Figure 5.25: Mean VOT values (in ms) for each of the monolingual English (left) and Arabic (right) children. N = 447.

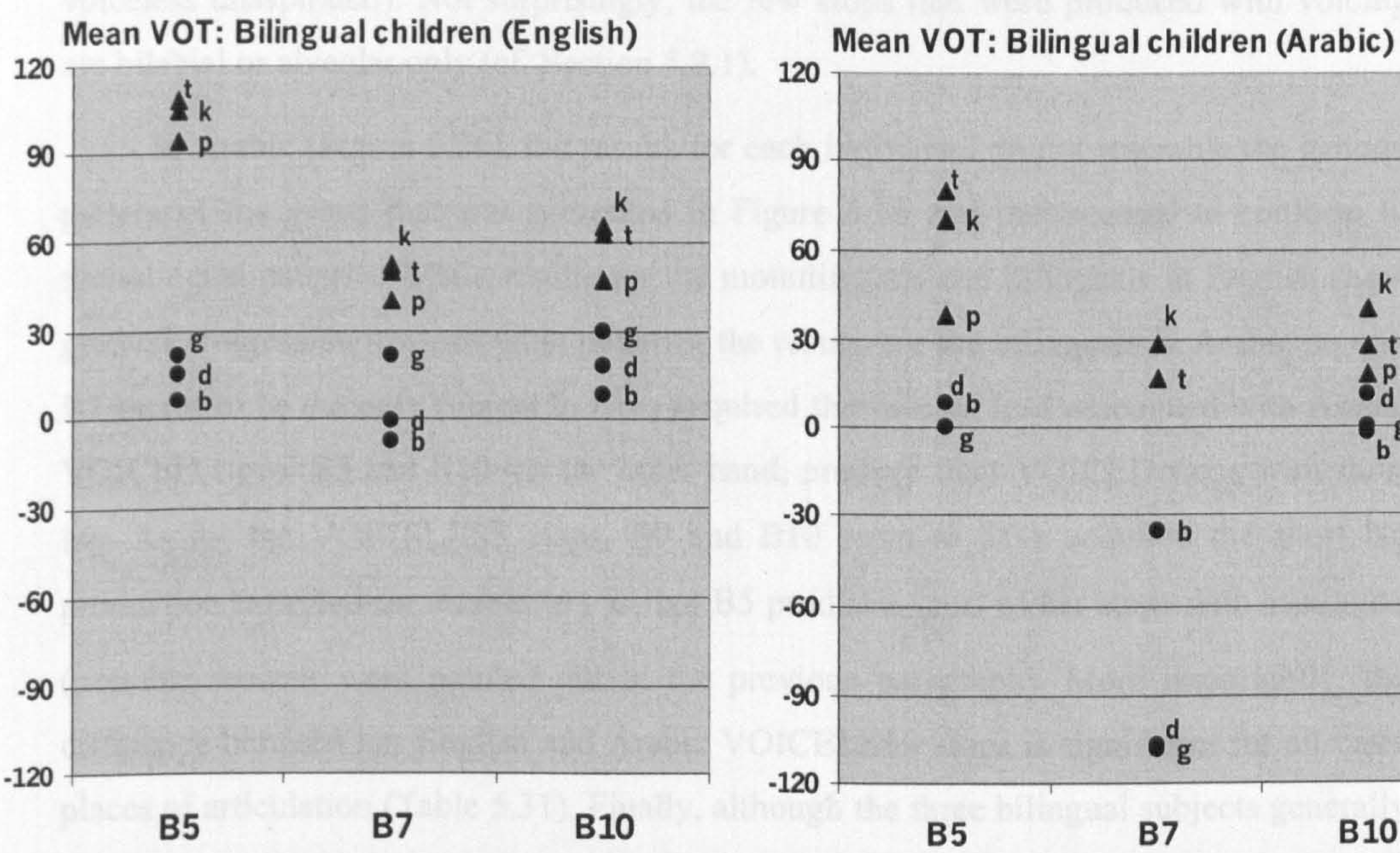


Figure 5.26: Mean VOT values (in ms) for each of the bilingual children in English (left) and Arabic (right). N = 415.

Moving on to the bilingual children's patterns in English (Figure 5.26, left), the results for each individual subject show expected patterns that are on the whole similar to monolingual ones. The VOICELESS stops are aspirated and the VOICED stops are produced with a short lag or voicing lead (for B7). Moreover, the measurements for /g/

and /p/ generally fit within the expected VOT continuum for English stops, which proves that the subjects have successfully acquired them though they are in principle absent from the inventory of Lebanese stops. Similarly to the monolingual groups, there are signs of gradual progression towards the adult model in that B10 has the most stable values for his VOT production with narrower ranges than those of the younger two subjects.

Individual differences include the fact that B5 has particularly long VOT values for her VOICELESS stops. This (i) may be due to her young age (ii) may be idiosyncratic in that her VOT values are longer in respect to both English and Arabic or (iii) may be a strategy used by B5 to keep her English and her Arabic VOICELESS stops separate, since her Arabic stops are aspirated as well. Other individual differences can be found in B7's results, as he generally has short VOT values compared to the other two bilingual subjects and is the only subject to produce a few of his VOICED English stops with voicing lead. With regards to his short VOT values, this may be an individual difference or may be due to the fact that B7's speech rate was slightly faster than that of the other subjects. As for voicing lead for his VOICED stops, this should not be surprising since VOICED stops in English are known to have two possible realisations (voiced and voiceless unaspirated). Not surprisingly, the few stops that were produced with voicing are bilabial or alveolar only (cf. Section 5.9.1).

In Arabic (Figure 5.26), the results for each individual do not resemble the general pattern of the group that was presented in Figure 5.16 and that seemed to conform to monolingual patterns. While results for the monolinguals and bilinguals in English show gradual progression towards adult patterns, the results for the bilinguals in Arabic do not. B7 seems to be the only subject to have acquired the voicing lead associated with Arabic VOICED stops. B5 and B10, on the other hand, produce their VOICED stops with short lag. As for the VOICELESS stops, B7 and B10 seem to have acquired the short lag production expected for Arabic /p t k/, but B5 produces most of her stops with aspiration (possible reasons were pointed out in the previous paragraph). More importantly, the difference between her English and Arabic VOICELESS stops is significant for all three places of articulation (Table 5.31). Finally, although the three bilingual subjects generally produce higher VOT values for their VOICELESS stops in Arabic than their VOICED stops in English, the difference is not as big as the one found between the monolingual English and Arabic subjects. This may suggest that the bilingual children are using other cues to achieve the perceptual difference between Arabic /p t k/ and English /b d g/.

Table 5.31: Summary of VOT means, standard deviations, and t-test results for stop production in English and Arabic by the bilingual and monolingual children. E = English; A = Arabic

		p		t		k		b		d		g	
		E	A	E	A	E	A	E	A	E	A	E	A
B5	Mean	94	38	108	80	105	70	6	8	15	8	22	0
	SD	30	25	50	40	32	24	8	11	13	9	12	0
	t-test	p = 0.003*		p = 0.080		p = 0.002*		p = 0.621		p = 0.246		p = 0.007*	
		p		t		k		b		d		g	
		E	A	E	A	E	A	E	A	E	A	E	A
E5/A5	Mean	52	23	75	30	62	40	5	-48	9	6	17	-2
	SD	12	20	24	15	22	7	8	46	9	8	10	41
	t-test	p = 0.003*		p = 0.000***		p = 0.001*		p = 0.000**		p = 0.376		p = 0.269	
		p		t		k		b		d		g	
		E	A	E	A	E	A	E	A	E	A	E	A
B7	Mean	41		50	17	52	28	-8	-35	-1	-107	22	-108
	SD	21		29	19	19	18	34	31	23	28	5	
	t-test			p = 0.000**		p = 0.02*		p = 0.041*		p = 0.000**			
		p		t		k		b		d		g	
		E	A	E	A	E	A	E	A	E	A	E	A
E7/A7	Mean	87	7	98	16	101	21	14	-32	25	-51	28	17
	SD	39	9	29	7	31	8	9	30	13	62	21	12
	t-test	p = 0.000***		p = 0.000***		p = 0.000***		p = 0.000***		p = 0.006*		p = 0.246	
		p		t		k		b		d		g	
		E	A	E	A	E	A	E	A	E	A	E	A
B10	Mean	48	18	63	28	66	41	9	-2	18	11	30	0
	SD	10	16	17	19	20	20	8	12	11	14	7	
	t-test	p = 0.000***		p = 0.000***		p = 0.005*		p = 0.005*		p = 0.141			
		p		t		k		b		d		g	
		E	A	E	A	E	A	E	A	E	A	E	A
E10/A10	Mean	52	12	59	21	70	27	5	-54	13	-42	17	-7
	SD	12	15	21	8	22	8	4	16	7	20	13	59
	t-test	p = 0.009*		p = 0.000***		p = 0.000***		p = 0.001*		p = 0.025*		p = 0.560	

5.9.6 Developmental changes for B7 and B10

A further examination of the differences in the production of Arabic VOICED stops by the two bilingual brothers B7 and B10 was conducted in order to try and find out how each of the patterns emerged in the production of each child and what factors affected their acquisition in such a way that only B7 acquired voicing lead in Arabic. A pilot study that was conducted 18 months before the current study made data collected from the two brothers available for analysis and comparison. At the time, only /b/ /d/ /t/ and /k/ tokens were collected for Arabic, but the VOT patterns obtained then offer the opportunity to trace the development of VOT patterns in the two subjects over a period of 18 months. Figures 5.27 and 5.28 juxtapose the VOT values obtained for each child at the time of each recording and for each language. Like the 2000 data, the 1998 data consisted of word-initial stops followed by a variety of vowels covering a wide range of the phonetic space.

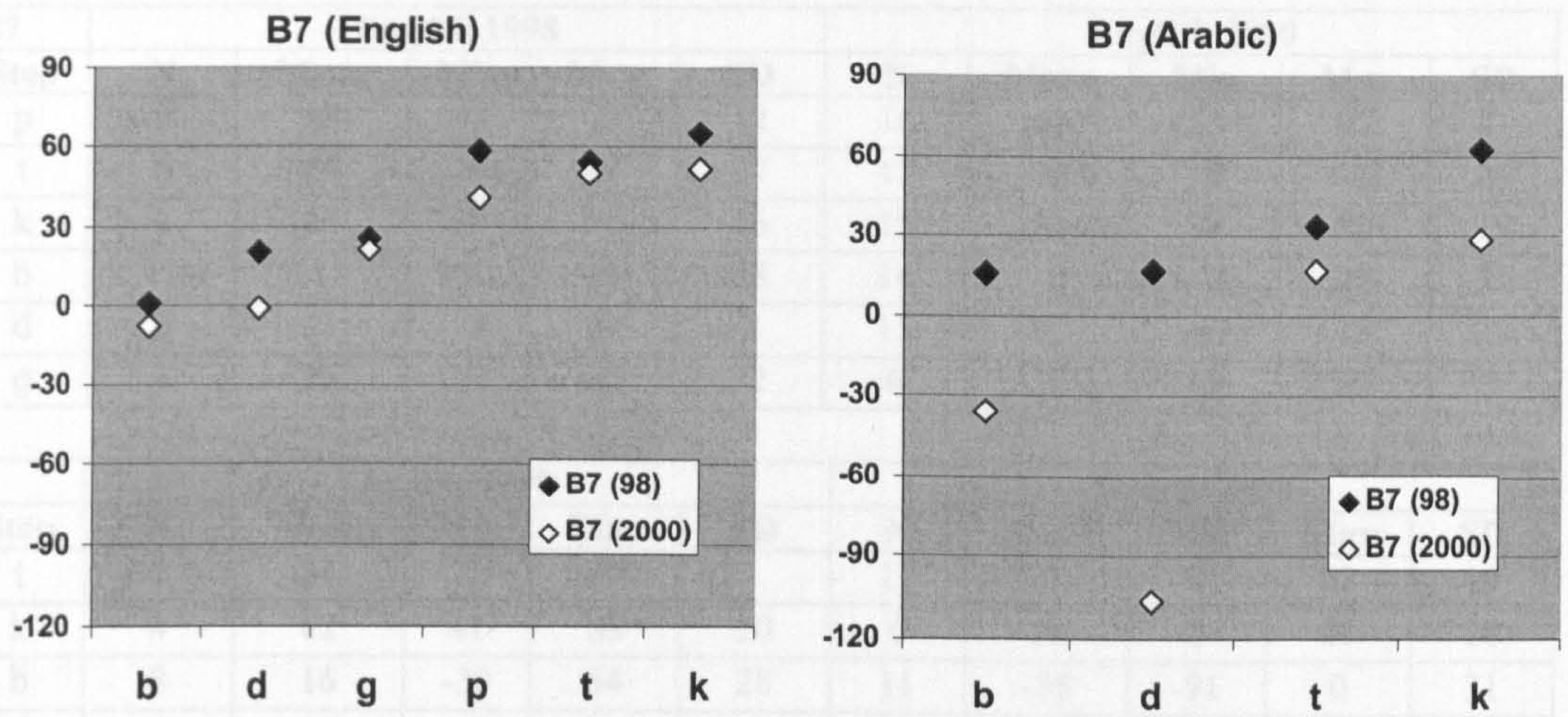


Figure 5.27: Mean VOT values for B7 in 1998 and 2000 in English (left) and Arabic (right). N = 176.

Looking at B7's developmental changes in English first, one can see that the VOT values for his VOICELESS stops have slightly been reduced since 1998, but that more change has affected his VOICED stops. It is interesting to note from the distribution in Table 5.32 that since the earlier recordings in 1998, B7 had already started using two patterns for his VOICED stops in English (voicing lead and short lag). The VOT changes in English are not very surprising considering the fact that younger children normally produce high VOT values that are later reduced with age (e.g. Foulkes, Docherty & Watt, 1999). A similar reduction is noted in B7's VOT measures for his VOICELESS stops in Arabic. While his stops were slightly aspirated in 1998, they are now produced with short lag, although the values from 1998 are actually closer to the values found for the adults in this study. Unlike the results for English, there is a significant difference between B7's Arabic /k/ values in 1998 and 2000 (Table 5.32). Still, the most surprising result for B7 is the difference in the VOT patterns for his VOICED stops between the two recordings. While in 1998 B7's /b d/ production in Arabic fell mainly in the short lag region and resembled the patterns currently observed for B5 and B10, his production seems to have changed noticeably by the time of the second recording. Most of his VOICED stops are now produced with a voicing lead, which is the pattern that is expected for Arabic but that is usually difficult to acquire. T-tests show significant difference between B7's VOT values for VOICED stops in 1998 and 2000 (Table 5.33). An investigation into the reasons behind this change as well as an interpretation of the results will be attempted in the following section.

Table 5.32: Detailed measurements of VOT means, ranges, and standard deviations in English and Arabic for B7 in 1998 and 2000.

B7	English 1998					English 2000				
Stop	**N**	**Mean**	**Min**	**Max**	**SD**	**N**	**Mean**	**Min**	**Max**	**SD**
p	9	**58**	42	73	12	16	**41**	15	92	21
t	6	**55**	29	97	27	17	**50**	16	122	29
k	6	**65**	39	91	16	15	**52**	19	80	19
b	11	**1**	-40	34	28	16	**-8**	-98	28	34
d	3	**21**	18	24	3	15	**-1**	-67	19	23
g	4	**25**	-21	49	32	6	**22**	15	28	5

	Arabic 1998					Arabic 2000				
Stop	**N**	**Mean**	**Min**	**Max**	**SD**	**N**	**Mean**	**Min**	**Max**	**SD**
t	7	**34**	23	69	17	17	**17**	0	62	19
k	4	**62**	41	89	20	6	**28**	0	48	18
b	8	**16**	-30	54	28	11	**-35**	-91	0	31
d	4	**17**	-17	33	16	5	**-107**	-152	-75	28

Table 5.33: T-test results comparing VOT distribution between B7's production in 1998 and 2000.

	/p/	/t/	/k/	/b/	/d/	/g/
English	p = 0. 010*	p = 0. 775	p = 0. 211	p = 0. 560	p = 0. 003*	p = 0. 842
Arabic		p = 0. 093	p = 0. 034*	p = 0. 005*	p = 0. 000**	

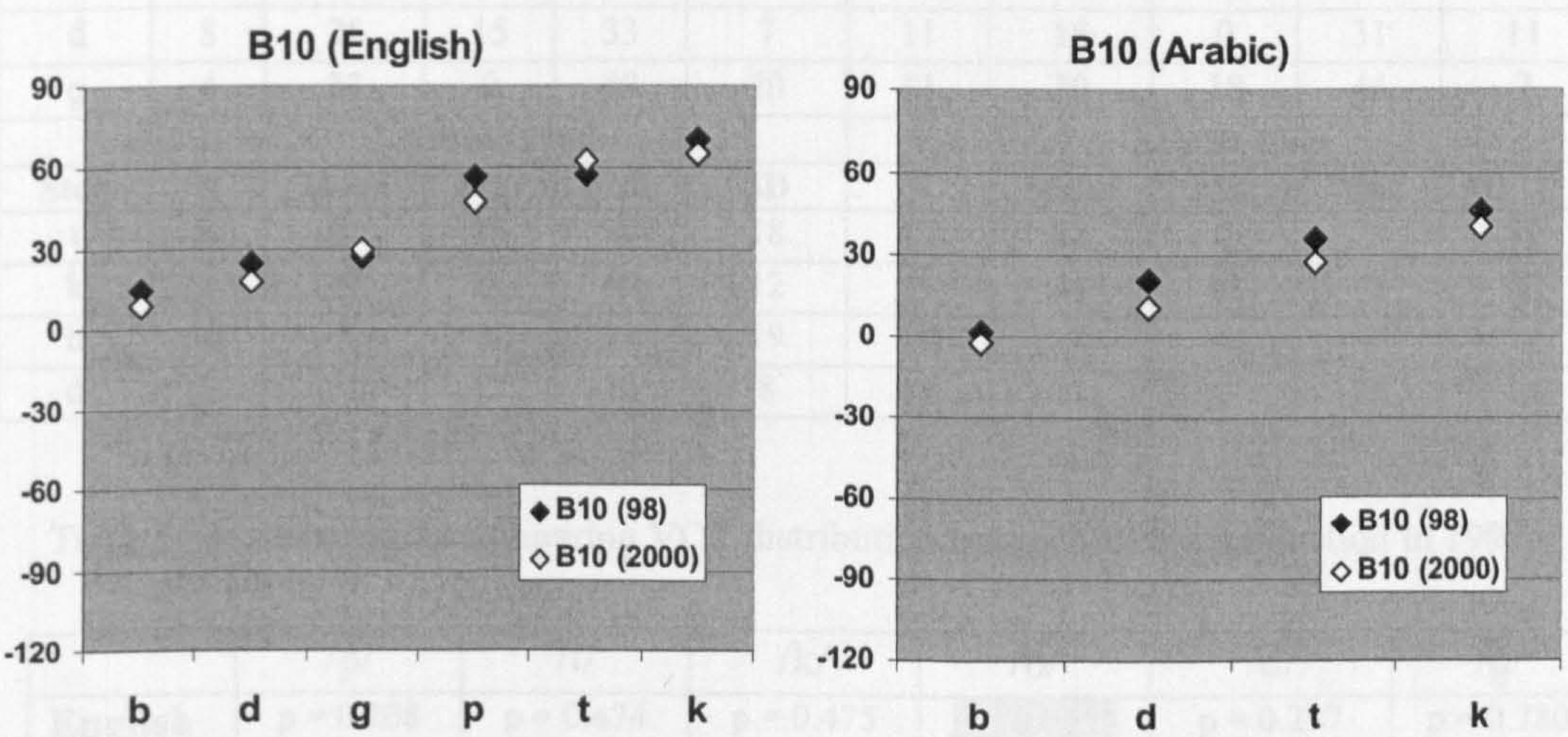


Figure 5.28: Mean VOT values for B10 in 1998 and 2000 in English (left) and Arabic (right). N = 192.

Looking at B10's developmental changes in English first (Figure 5.28 and Table 5.34), one can see that, like B7, the VOT values for both his VOICED and VOICELESS stops have generally been reduced, though to a lesser extent in B10's case and not for all the stops. There is actually no significant difference in B10's English VOT production between 1998 and 2000 (Table 5.35). Such a difference between the changes for the two brothers is expected considering the fact that VOT patterns for the younger subject will be developing faster and greater while the older subject will have developed more stable patterns that are closer to the adult model. In fact, B10's VOT for VOICELESS stops was significantly different across the two languages at the time of the early recordings in 1998 ($p = 0.025$ for /k/ and $p = 0.01$ for /t/), while B7 only achieved that in the later recordings in 2000 (Section 5.9.4.2). However, the same cannot be said about B10's VOT development in Arabic, because his VOT patterns for the VOICED stops are still predominantly produced with short lag and do not resemble the adult model, though the patterns for his younger brother do. Note that, like in English, B10's VOT values in Arabic have dropped for both VOICED and VOICELESS stops but not significantly (Table 5.35). Also, while the values for VOICELESS stops are now closer to the expected short lag pattern, the values for VOICED stops have only been slightly reduced.

Table 5.34: Detailed measurements of VOT means, ranges, and standard deviations in English and Arabic for B10 in 1998 and 2000.

B10	**English 1998**					**English 2000**				
Stop	**N**	**Mean**	**Min**	**Max**	**SD**	**N**	**Mean**	**Min**	**Max**	**SD**
p	16	**57**	35	98	20	17	**48**	34	65	10
t	6	**58**	44	84	11	16	**63**	42	107	17
k	4	**71**	60	80	8	13	**66**	40	101	20
b	12	**15**	0	24	5	16	**9**	0	27	8
d	8	**25**	15	33	7	11	**18**	0	31	11
g	4	**27**	0	48	20	11	**30**	18	44	7
	Arabic 1998					**Arabic 2000**				
Stop	**N**	**Mean**	**Min**	**Max**	**SD**	**N**	**Mean**	**Min**	**Max**	**SD**
t	8	**37**	15	59	18	17	**28**	0	79	19
k	3	**47**	36	60	12	11	**41**	18	72	20
b	10	**2**	-30	21	19	16	**-2**	-44	8	12
d	3	**21**	17	30	8	17	**11**	0	36	14

Table 5.35: t-test results comparing VOT distribution between B10's production in 1998 and 2000.

	/p/	/t/	/k/	/b/	/d/	/g/
English	p = 0.208	p = 0.474	p = 0.475	p = 0.045*	p = 0.257	p = 0.780
Arabic		p = 0.396	p = 0.548	p = 0.614	p = 0.138	

5.9.7 Code-switched tokens

So far the discussion has concentrated on either the English tokens produced by the bilingual children during the English sessions or the Arabic tokens produced during the Arabic sessions. As was found in the previous two chapters, the English code-switches that were produced by the children during the Arabic sessions displayed phonetic patterns that were different from the patterns observed during the English-only sessions. All the English word-initial stops that were produced during the Arabic sessions were therefore extracted for analysis in order to examine whether the VOT patterns in these tokens conformed with the patterns produced by the children during the English-only sessions. While most tokens were taken from isolated words (e.g. 5.1), others were part of an utterance (e.g. 5.2).

(5.1) Mother (pointing at a kettle): [ʃu haɪda]?
What that (masc.)?
'What is that?'
Child: [ˈkɛtəl]
<u>KETTLE</u>

(5.2) Child (describing a an action): [nɑtˤtˤo bɪl pu:l]
jump-past-3rd pers.pl. in <u>POOL</u>
'they jumped in the pool'

The VOT tokens were not numerous but still needed to be examined separately from the data presented so far in order to test the difference in patterns that the children exhibited in the previous chapters with regards to /l/ and /r/ production. As most of the words are the same ones that the bilinguals produced in the English sessions, this allowed direct comparison of two productions of the 'same' word. Tables 5.36 shows VOT measurements of the English tokens produced by the children in the Arabic sessions and, where possible, in the English sessions.

Table 5.36: VOT measurements of English target words produced by B5 during the Arabic and the English sessions

B5	Gloss	Arabic sessions		English sessions	
		IPA	VOT (ms)	IPA	VOT (ms)
p	pool	pʰʉ̣əl	99	pʰʉʊ	108
	pepper	ˈpʰɛpʰəɾ	112	ˈpʰɛpə	51
	pink	pʰ̃iŋk	96-118	pʰ̃iŋk	109
	painting	ˈpʰẽ·nt͡ɪn	36	ˈpʰẽ·ntʰ̃ɪng ˈpʰẽ·ntʰ̃ɪng	109 99
	peeling	ˈpʰi::lɪŋ	105	ˈpʰi:lɪŋg	90
t	tent	tʰɛnt	55	tʰɛnt	103
	Teletubbies	ˈtejtʌbi:	38	ˈtʰɛlɪtʌbi:	48

	tummy	ˈtʰʌmi	80	ˈtʰʌmi	91
k	coffee	ˈkʰɔfiː	95-94-93	ˈkʰɔfiː	76-85
	can	kʰanː	124	kʰanː	70
	car	kʰɑɾ	90	kʰɑː	136-102
	kettle	ˈkɛtəl	37	ˈkʰɛtəɫ	70
	castle	ˈkʰa̯sədə	96	ˈkʰasɫ̩	93
b	bottle	ˈbɔtəl	0	ˈbɔtəɫ	12
	butterfly	b̥ʌtəˈfːaɪ	34	b̥ətəˈflaɪ	22
g	garlic	ˈgɑːlɪk	31	ˈgɑːlɪk	36
	goat	goːt	25	gəʊt	21
	garden	ˈgɑːrtən	64	ˈgɑːdən ˈgɑːd̥ən	21 59

B7	**Gloss**	**Arabic sessions**		**English sessions**	
		IPA	**VOT (ms)**	**IPA**	**VOT (ms)**
p	po	pʰoᵘ	56	pʰɔə	92
	purple	ˈpɜːpəl	29	ˈpɜːpɵ	22
	picture(s)	ˈpɪktʃɜːːz̥	0	ˈpʰɪktʃə	30
t	teapot	ˈtiˑpot	13	ˈtʰiˑpɔt ˈtʰiˑpɔtʰ tiˈpɔt	42 39 35
	tummy	ˈtʊmiː:	12	ˈtʰʊmɪ	42
	tinky	t̃ɪŋkɪ	15	t̃ɪŋkɪ	25
k	cow	ka̯ʊ	23	kʰaʊ	53
	carrot	ˈkɐɾ̥ət	25	ˈkʰaɹət	57
	coffee	kʰɔˈfiː	34	ˈkʰɔfi̥	40
	can	kʰanː	67	kʰanː	79
b	bottle	ˈbɔtəl	-54	ˈbɒtɐ	-67
	beer	bːiəɾ̥	-165	bɪə	0
	bed	bɛd	8	bɛd	0
	butter	ɓʌˈtʰər	- 60	ˈɓʌtʰə	- 31
	butterfly	bətəɾ̥ˈflaɪ	0	ˈbʌtᵊflaɪ	12
d	duck	dʊk	12	ndɜks d̥ɜk dɜk	-97 21 6
	Dipsy	ˈdɪpsi ˈd̥ɪpsiː	17 7	ˈd̥ɪpsiː	18
g	guitar garden	giˈtɑːːɛlɾi ˈgɑɾ̥dən	-108 20	gɛnˈtɑ̯ ˈgɑːdən	28 24

B10	**Gloss**	**Arabic sessions**		**English sessions**	
		IPA	**VOT (ms)**	**IPA**	**VOT (ms)**
p	po	pʰəʊ	95	pʰəːːʊ	166 (slow)
	pool	puːᵊl	28	pʰʉɤ	84
	peg	pʰɛg	40	pʰɛg	65

t	teapot	ˈti·pɔt˺	19	ˈtʰiːpɔtʰ	76
	tinky	ˈtɪŋki	26	ˈtʰɪŋkʰi	44
	Teletubbies	tɛlɪˈtʌbis	13	ˈtʰɛlɪtʌiːz̥	59
k	castle	ˈkasl̩	30	ˈkʰasɵ	90
	cup	kʊp’	22	kʰʊp	61
	cartoons	kʰa̹˞ˈtʰɜːn	31	ˈkʰɑːtʰʉːnz̥	69
	kitchen	kʰɪˈtʃɪn	46	ˈkʰɪʔtʃən	46
b	back	bɛ·k	9	bakʰ	3
	bugs	bʌgz̥	9	bœ̹gz̥ bœ̹gz̥	6 9
	beer	biɾ̹	0	bɪə	9
d	Dipsy	dɪpˈsiː	16	d̥ɪpˈsiː	31
	deer	diːɾ		dɪə	31
	guitar	giˈtɑ̹·ɾ̥	13	gɪˈtʰɑː	34

Looking at B5’s results first, it seems that the English production patterns prevail in her productions during both the Arabic and the English sessions. The VOICELESS stops are highly aspirated, while the VOICED stops are produced with short lag (apart from one voiceless production of ‘garden’). In cases where there was a noticeable difference between the productions in the two languages (highlighted in grey), the pattern was not always predictable, i.e. the VOT values in the code-switched items did not necessarily follow the Arabic pattern in the same way other features do. For instance, while ‘painting’, ‘tent, and ‘kettle’ have shorter VOT values when produced during the Arabic sessions as opposed to the English ones, ‘pepper’ and ‘can’ follow the opposite pattern. Therefore it seems that while B5 has manipulated several features in the code-switched tokens that made their production qualitatively different from comparable productions during the English sessions, VOT is not manipulated in the same way. However, there are two important things to remember: first, even in the main data analysed in Section 5.9, it was found that B5 produced high VOT values for her VOICELESS stops in both languages (though the difference is significant); second, it was also found that B5 produced both her English and her Arabic VOICED stops with voicing lag, so no change in the production of VOICED stops would have been expected for the code-switches.

As for B7 and B10, the differences in their VOT production between the Arabic and the English sessions are not random. Instead, they follow a much more consistent pattern in that their VOICELESS stops often have shorter VOT values in the Arabic than in the English productions of the ‘same’ words (highlighted in grey). This suggests that the two bilinguals can manipulate VOT features in the same way that they manipulate other vocalic and consonantal features to make the words sound more Arab- or English-like.

5.10 Summary and discussion

5.10.1 Discussion with relation to the aims of the study

An attempt will now be made to answer the four questions that gave rise to the current study (Section 5.7). The first two questions will be answered together as they are related. These are:

1 Do English-Arabic bilinguals acquire separate VOT patterns for each of their languages?

The results obtained from this study do offer support to the view that bilingual subjects acquire separate production strategies for their stops in each language. Each of the three bilingual subjects appear to have distinct patterns for English and Arabic, but achieve the distinction by various means. For instance, B5 generally has aspirated VOT values for her VOICELESS stops in both languages, but still produced significantly higher values in English than in Arabic. B7, on the other hand, had short VOT values for his English VOICELESS stops, but he still produced significantly shorter values for Arabic.

2 Are their patterns of production in each language similar to those of the monolingual controls in the study?

While the patterns observed for all three bilingual children in English follow the monolingual model, the patterns in Arabic are not the same for all subjects, especially with respect to VOICED stops.

In English, the monolingual and bilingual children produce their VOICELESS stops with long lag, and their VOICED stops with short lag, with more overlap between the two categories of stops than was found for the adults. The younger subjects (E5 and B5) generally produce longer average VOT than the older subjects, and exhibit greater variability in terms of wider ranges and larger standard deviations. The oldest subjects (E10 and B10) have more stable VOT patterns than the younger ones, and their patterns are closer to what is known about the adult model, especially in terms of the narrower ranges they use.

Individual differences include the fact that some children produce particularly long VOT values (B5 and E7), while others produce particularly short ones (B7). Moreover, B7 is the only child in the group who uses two patterns for his VOICED stops in English, long lead and short lag, but his prevoiced stops are fewer than the short lag ones and are either accompanied by features that are characteristic of nasal sounds similar to the ones found for his mother (BF7), or that are realised as implosives.

In Arabic, the monolingual subjects show signs of acquisition of the voicing lead/short lag pattern for VOICED and VOICELESS stops, and a potential developmental pattern for the acquisition of voicing lead, as only the oldest subject (A10) has acquired it for the three places of articulation. As for the bilingual subjects, only one out of three children (B7) seems to have acquired voicing lead, and it is not the oldest child.

The results in this study are similar to those by Deuchar & Clark (1995) for their English-Spanish subject, though their study was conducted on a younger subject (age 1;7 till 2;3). They also found short lag as a realisation of Spanish VOICED stops by the child, and explained this with reference to the parental model, which also showed short lag. They further predicted that a fairly adult-like model would develop at around the age of seven. Similar results were found by Konefal & Fokes (1981), whose four-year-old subject produced her Spanish VOICED stops with short lag, while her seven-year-old sister had acquired voicing lead. However, in the current study, age does not seem to be the only factor for the acquisition of such a complex feature by the bilingual subjects, as the adult-like model has started to develop in the seven-year-old subject's patterns but not the ten-year-old one. Moreover, an examination of B7's prevoiced tokens reveals his use of devices that might be aiding the production of voicing lead, such as nasal-like features preceding the stop or implosives, but that were also found in his mother's speech (BF7).

Note that B7's use of a nasal-like sound is not dissimilar to that of Allen's (1985) monolingual French subjects who preceded French VOICED targets with a nasal or a vowel segment that permitted continuous voicing, Macken & Barton's (1980) monolingual Spanish subjects who spirantised their Spanish VOICED stops, or Heselwood & McChrystal's (2000) Panjabi-English subjects who showed both features. In all four cases (Arabic, French, Spanish and Panjabi), the subjects are choosing an articulation that does not involve a complete obstruction of the airstream in order for them to prolong the voicing articulation. In B7's case, there is the added reason that pre-nasalisation of fully voiced stops was also found in his mother's speech. B7's production of implosives can also be considered as another strategy used to maintain transglottal pressure, and hence glottal vibration (Heselwood, 1998).

The difference between the two brothers B7 and B10 is surprising when considering the fact that the ten-year-old child might be expected to have greater command over his stop articulations and more control over glottal and supraglottal events than the seven-year-old child. A possible explanation will now be attempted. In Section 5.9.5, an investigation into the developmental changes in VOT for the two brothers B7 and B10 showed that B7 had only recently acquired the production of voicing lead for his VOICED stops in Arabic. Data from recordings made 18 months prior to the main corpus of recordings show that B7 had a similar pattern to B5 and B10 with regards to the production of VOICED stops. Two main events had occurred in the bilingual brothers' lives since the first recordings were made in 1998 and may have affected their linguistic behaviour by the time the second set of recordings was made in 2000. First, the two brothers started attending a weekend Arabic school where pupils learn reading and writing in Arabic and receive Arabic input from teachers and other pupils. Second, their parents made friends with a Lebanese family that had recently moved to a nearby city and had been socialising with them since.

The outcome of these changes was an increased input in Arabic for the two brothers, which was lacking at the time of the first recording when the parents were the main source of Arabic to the children. Since such increase in Arabic input normally comprises an increase in the amount of voicing lead that might become more salient in the brothers' environment, it might have affected B7's acquisition of this feature though not B10's. Knowing that the two brothers have experienced the same changes in Arabic input, speculations about the uneven change in their behaviour suggest that it can mainly be attributed to their age. While B7 was still five years old when he experienced the increased input in Arabic, his brother was eight and was probably past the 'critical' age required for the acquisition of such a complex feature (Flege, 1995). Like in monolingual situations, the acquisition of certain complex features that require early and extensive exposure might therefore be delayed or not acquired if these features are lacking in the input that the bilinguals receive.

B7's acquisition of voicing lead proves that, when the two conditions of input and age are met, bilinguals will follow similar acquisitional patterns to those of monolinguals, or even 'catch up' with them. In the process of doing so, bilinguals will exhibit developmental patterns that are once more similar to monolingual ones, such as the use of short lag instead of voicing lead (Allen, 1985; Macken & Barton, 1979), or the use of continuants preceding the stops in an attempt to acquire voicing lead.

What is important to consider, however, is that B10's use of short lag is not necessarily caused by an influence from English, but is possibly due to the fact that B10 did not receive enough input in Arabic voicing lead at an early age for him to master the complex articulatory features required for its production. Similarly, B7's use of strategies that are similar to those used by monolinguals points to the necessity of looking for normal developmental processes in order to explain the speech patterns observed in bilinguals before resorting to explanations based on language interaction and interference. Such processes may be due to the articulatory difficulty associated with some gestures (in this case consonantal ones) that can delay children's acquisition of speech timing, regardless of whether they are monolingual or bilingual.

3 Are there signs of influence from one language onto the other in the bilinguals' production and what are the factors that affect such influence?

As mentioned in the previous two chapters, there are two different types of influence between the two languages that need to be examined separately. On the one hand the slightly more aspirated VOT values for Arabic VOICELESS stops that were found for B5 compared with her monolingual counterpart and the incomplete acquisition of voicing lead by two of the bilinguals suggests influence from their dominant language. However, due to the fact that the experiment is conducted on children, it is important to be able to distinguish between features that the bilingual subjects have not acquired because of their bilingual background and those that are missing due to the fact that their languages are still developing. For instance, although B5's VOICED stops in Arabic are produced with short lag rather than long lead, such patterns cannot be solely attributed to her English dominance due to the fact that A5 shows more or less similar features. By contrast, we can conclude that B10 has not acquired the pattern for Arabic VOICED stops, since he still produces short lag. On the other hand, A10, his monolingual counterpart shows signs of having mastered the production of voicing lead for all places of articulation.

The second type of influence concerns the code-switched data. The language context in which the children produced target VOT tokens turned out to be crucial for the interpretation of the resulting production patterns. Very few studies on VOT have specified the linguistic context in which the bilingual subjects' production occurred and the effect this may have had on the resulting patterns. In this study B7 and B10 show strong awareness of the language context and of their ability to manipulate VOT patterns of English words depending on whether these were produced in an Arabic or English context. If the patterns found for the English words produced during the Arabic sessions were to be included with the rest of the data that were analysed in Section 5.9, one might have reached the erroneous conclusion that the bilinguals have not acquired separate patterns for each of their languages.

4 Are the patterns for the monolingual subjects in this study similar to the ones normally described in the literature and therefore expected for each language?

a. Children

With regards to the children, this question is difficult to answer in the case of Arabic, since there are no VOT accounts for Arabic-speaking children. As for the English results, these seem fairly similar to the studies reviewed in Section 5.5.2, which suggest that children produce VOT with longer duration and more variability than adults do, and that adult-like consistency is usually achieved when reductions in the duration of speech sounds and in variability gradually take place as children become older and until approximately 10-12 years of age. Therefore, an important factor that needs to be accounted for when analysing child speech is the fact that results obtained for a given child often reflect developmental stages rather than hard-wired production patterns. For instance, the results for the oldest monolingual subjects in this study are the closest to what is known about VOT patterns in the subjects' respective adult communities, while the results for the younger subjects vary and show incomplete acquisition in many respects.

One important result is that of the VOT pattern found for VOICED Arabic stops by the monolingual five- and seven-year-olds. Similar to results obtained in monolingual (Allen, 1985; Macken & Burton, 1980) and bilingual studies (Deuchar & Clark, 1995; Heselwood, & McChrystal, 2000; Konefal & Fokes, 1981), results from this study show that some adult patterns develop later than others. In this case, the complexity of the articulatory gestures involved in the production of voicing lead and the difficulty of co-ordination of laryngeal control with a particular supralaryngeal articulatory gesture may delay the children's acquisition of voicing lead and prevent them from mastering it at an early age (Kewley-Port & Preston, 1974; Zlatin & Koenigsknecht, 1976). In physiological terms, the difficulty in producing voicing lead is due to the fact that, when the pressure increase behind the stricture reaches the level of subglottal pressure, transglottal airflow ceases and voicing is impossible (Ohala, 1997: 687). Children, with shorter vocal tracts, will be unable to sustain voicing in this condition for as long as adults, which may prompt them to seek compensatory strategies if they are attempting to match the values of adult speech. Some of these strategies include spirantisation of voiced stops, which was observed by Macken & Barton (1980) for a 4-year-old Spanish-speaking child; Allen (1985) reported prenasalisation and prevocalisation with an oral or nasalised vowel by 1;9-2;8 aged French speakers, while Heselwood & McChrystal (2000) reported all of the above features in their 10-year-old Panjabi-English bilinguals, with different subjects choosing different strategies.

Therefore, while adult-like VOT patterns in English normally start appearing around the age of two, it must be kept in mind that children who acquire languages that contrast voicing lead with voicing lag might develop those patterns at a later age.

Similarly, results obtained for both the monolingual and bilingual children in this study show greater amount of overlap between the VOT values for English VOICED stops and Arabic VOICELESS stops compared with the findings for the adults. There was no significant difference between the distributions for English /b d g/ and Arabic /p t k/

for any of the age groups seven or ten, mainly due to the fact that the children produced shorter VOT values for Arabic VOICELESS stops than those of the adults. The children's results support the view that the phonological voicing feature of a stop cannot be made on the basis of VOT alone (Docherty, 1992: 116; Klatt, 1975: 695). Such an observation is also very important when judgements are being made about the bilinguals' ability to separate between the two sets of stops in either language.

b. adults

Moving on to the adults, the overall values obtained for the parents in this study follow the expected pattern for each group of speakers. In English, the monolinguals' parents mainly produce short lag with some voicing lead for their VOICED stops and long lag for their VOICELESS stops, while the bilinguals' parents apply the Arabic VOT patterns on their VOT production by mainly producing voicing lead for their VOICED stops and short lag to slight aspiration for their VOICELESS stops.

Interesting observations include the fact that most speakers keep the distributions for their VOICED and VOICELESS stops quite separate (cf. Docherty, 1992). Moreover, the small number of the VOICED stops that were produced by the monolingual English parents with voicing lead confirms the fact that descriptions of the phonetic realisations of English VOICED stops should not be restricted to one single short lag category (e.g. Docherty, 1992; Lisker & Abramson's study; 1967). The study also offers further support to observations made by Docherty (1992) about exceptions to the apparently universal rule of place of articulation effect on VOT in VOICELESS stops, as some of the speakers had higher VOT means for bilabial than for alveolar (EF10) or even velar (EM5) stops.

The VOT patterns found for /p/ and /g/ as produced by the bilinguals' parents confirm the fact that they have acquired the production of these two sounds despite their rare occurrence in their native language, but have applied the Arabic phonetic implementation for VOICELESS sounds by producing /p/ with short lag to slight aspiration and for VOICED sounds by producing /g/ with voicing lead. However, one of the bilinguals' parents (BF7) frequently produced audible and acoustically detectable nasals and/or vowels before her VOICED stops in English. It was surprising to find these features in the production of one of the adults, since, as was mentioned in the previous section, they are normally noted in the speech of children who are still in the process of acquiring voicing lead.

As for the Arabic results, the Lebanese adults in this study have similar VOT patterns to those found for some of the other Arabic dialects (Section 5.5.3) and confirm the fact that the overwhelming VOT pattern for Arabic is that of voicing lead for VOICED stops and short lag to slight aspiration for VOICELESS stops. Although there is overlap between the VOT distributions for Arabic /p t k/ and English /b d g/, they seem to occupy different ranges along the VOT continuum (as proposed in Figure 5.2). The difference between the distributions for the two categories of stops was not significant for all subjects or for all places of articulation, but nevertheless supports Cho & Ladefoged's (1999) and Docherty's (1992) views that the boundary that separates between unaspirated and aspirated stops as suggested in the literature is arbitrary. This issue will be discussed in more detail in the following section.

Overall, the patterns for the monolinguals' parents and the bilinguals' parents look similar, which is also expected since Arabic is the L1 for the bilinguals' parents. There was a tendency for the bilinguals' parents to produce longer VOT values for both VOICED and VOICELESS stops than the monolinguals' parents, but the difference was not significant for all speakers or all places of articulation.

Despite the extended exposure to English by the bilinguals' parents for a period of 10 to 15 years, it seems that the subjects are still applying their native VOT patterns onto their stop productions in both English and Arabic, as most of the distributions showed no difference between the English and Arabic stops produced by these speakers. These results are surprising when compared with ones in other studies that have found significant long term changes in adult bilinguals' production in their L1 and/or L2 following extended exposure to the L2. With respect to changes to the L1, Flege (1987) examined VOT production in native speakers of American English who had lived in Paris for more than 12 years and native speakers of French who had lived in Chicago for a similar period. When compared to groups of monolingual controls, the American English speakers living in Paris had developed shorter VOT values in English than the American controls, while the French speakers living in Chicago had developed longer VOT values in French than the monolingual controls. Similar evidence was found by Major (1992).

More interestingly, a study by Sancier & Fowler (1997) also found evidence for short-term production changes in adult bilinguals, along with signs of both the L1 and the L2 of speakers being affected following periods of exposure to either language. In their study, an adult native speaker of Brazilian Portuguese who had learned English from the age of 15 and had been living in the United Sates for four years showed evidence of change in VOT production in both her languages depending on whether she was taped following several months' stay in Brazil or after she had been in the US for several months. After visits to Brazil, VOT for her VOICELESS stops in both English and Portuguese was significantly shorter than after several months' stay in the US.

No such effects seem to be operating in the bilinguals' parents in this study, but some observations could be offered to explain the disparity in the results. While all the bilingual's parents in this study are native speakers of Lebanese Arabic, the subjects in Flege's (1987) and Major's (1992) studies were married to native speakers of their L2 (e.g. the American subjects in Flege (1987) were married to French speakers while the French speakers were married to American speakers). This may have affected the results due to the added exposure to the L2 that the subjects in those two studies will have had at home. As for Sancier & Fowler's (1997) subject, she is reported as having actively studied English intensively before and after she moved to the US, and as having reached an advanced level of proficiency in English at the time the study started. The bilinguals' parents in this study are (self-reported) intermediate speakers of English and did not have any intensive training in English before or after they arrived in the UK.

5.10.2 Voicing timing in models of speech production

At this point, it is important to re-examine the three VOT categories (voicing lead, short lag, long lag) that have been used throughout the study as labels for the VOT values that were found for English and Arabic stops. As was noticed at several stages of the data analysis (Section 5.9), though English VOICED stops and Arabic VOICELESS stops have often been described in the literature as being produced with short lag, most the VOT values that were found for English /b d g/ for the adults and some of the children in this study were shorter than the values found for Arabic /p t k/. It is difficult to draw a line between the two sets of values due to the considerable overlap between the two distributions, but the results certainly suggest that different languages may be oriented to different portions of the VOT spectrum than the three categories would predict.

Though Cho & Ladefoged (1999) note that four phonetic categories could be adopted instead of three, they later backtrack on their idea when it comes to presenting these categories in a model of voicing timing, and return to the three-category choice due to their main interest in having enough categories to account for the number of contrasts in languages. However, one must not forget that there is very detailed phonetic learning that must take place in a child acquiring English or Arabic (or both), a kind of detail that is not necessarily oriented towards the acquisition of contrasts and that, in the case of voicing timing, is more specific than what the three possible categories suggest. This issue has been discussed extensively by Docherty (1992: 90), who criticises the overwhelming interest in accounting for contrast both within- and between-languages in models of voicing timing at the expense of detailed phonetic characteristics of languages or accents. Some of these models will now be reviewed.

Voicing contrast has been a major topic in phonology and phonetics over the last few decades. As already noted, the term 'voicing' can be understood in two ways: (i) physiologically, as the presence of vocal fold vibration during the closure phase of a given consonant; (ii) phonologically, as the abstract distinctive feature [±voice] which has a number of acoustic and articulatory correlates, only one of which is the presence or absence of vocal fold vibration. Several models have been proposed in order to account for the mapping of phonological categories relating to the voicing contrast onto the level of phonetic realisation of voicing timing. These include models that are feature-based (Chomsky & Halle, 1968), segment-based (Keating, 1984), function-based (Kohler, 1984), gesture-based (Goldstein & Browman, 1986), and parameter-based (Docherty, 1992; Cho & Ladefoged, 1999). Docherty (1992) provides a comprehensive review of most of these models and discusses the general problems related to them. Three of the models will be reviewed briefly in this section: these are Keating's (1984) and Cho & Ladefoged's (1999), as they concentrate on VOT in stops and are based on large-scale cross-linguistic studies, and Docherty's (1992), as it is based on data from Standard British English and provides a step towards a better representation of voicing timing in a model of speech production.

Keating's (1984) segment-based model consists of a rigidly structured view of the relation between the phonological feature [±voice] and its specific phonetic implementation. She proposes that binary phonological feature values (her model is only applicable to languages with a 2-way contrast) can be implemented as categories chosen from a fixed universally specified set consisting of 3 categories: fully voiced, voiceless unaspirated, and voiceless aspirated. The number of categories is chosen following the ultimate number of contrasting phonetic types in languages and corresponds to the standard division of the VOT continuum into voicing lead, voiceless unaspirated, and voiceless aspirated in initial position. These phonetic categories are further realised as articulatory and acoustic parameters represented continuously in time, although Keating does not elaborate much on how this is done (Figure 5.29).

Keating (1984) assumes that surface phonetic variation, within and across languages, may derive in a synchronic grammar from the interaction of three relatively simple systems: (i) the possible phonological features and their values; (ii) the possible phonetic category mappings; (iii) the phonetic detail rules accounting for variation within these phonetic categories. She considers that the rules of phonetic category implementation are language-specific and draw on the universal set of phonetic categories; and furthermore that the observed variation in the VOT distribution between languages can be derived by a general principle of 'polarisation' of adjacent categories along the voicing dimension. The notion of polarisation implies that, within the limits of implementation chosen for a given language, i.e. the phonetic categories, there is maximal separation of the distribution of values in order to allow for a robust contrast; for instance, {voiceless unaspirated} could polarise to a low range in a language like English that contrasts {voiceless unaspirated} with {voiceless aspirated}, and to a high range in a language like Polish which contrasts {voiced} with {voiceless unaspirated} (Keating, 1984: 48).

Phonological features ⟶ [±voice]
↓
Phonetic categories ⟶ {voiced, vl. unasp., vl. asp.}
↓
Phonetic realisation ⟶ Polarisation and specific quantitative rules.

Figure 5.29: Summary of Keating's (1984) model of phonological representation of voicing contrast in stops.

In sum, Keating (1984) tries to show that certain cross-language phonetic differences can best be expressed as differences in the realisation or implementation of phonological feature contrasts as phonetic categories. The three categories express the maximum number of contrasts found along the voicing dimension, although there are subtle differences in the exact position where the categories lie, which can be accounted for by principles such as polarisation.

Although Keating's model has been adopted by many researchers working on cross-linguistic comparison of the stop voicing contrast, it has obvious limitations. First, even for languages with the same syllable-initial contrast, for example {voiceless unaspirated} and {voiceless aspirated} in both English and German, some difference in VOT between the two languages can be found for the same phonetic category in comparable data sets and cannot be clearly explained by polarisation (e.g. Whitworth, 2000). Second, the model overemphasises the role of VOT in the realisation of the voicing contrast, with less consideration being given to other articulatory and acoustic correlates that play a role in the stop voicing distinction, such as the closure duration of the stop, the amplitude of the burst, etc. Third, the categories proposed by Keating are too rigid and abstract. On the one hand, as Docherty (1992: 12) notes, there are no clear boundaries between aspirated and unaspirated stops due to the fact that the delay in voice onset is more of a continuous process. On the other hand, within the phonetic categories resulting from the contrast, there are subtle language-specific differences that are not accounted for by a binary segmental representation, mainly with respect to fine temporal aspects of the phonetic realisation (Docherty, 1992: 74).

Docherty's (1992) model emerged following his own quantitative study of voicing timing in obstruents in Southern British English in a range of environments. One of the major findings from the study was the large degree of overlap between VOT for the VOICED and VOICELESS categories in the subjects' production. The values obtained varied systematically across speakers and contexts and were too detailed to fit into a model with three rigid phonetic categories. Docherty (1992: 89) noted that while most available models of voicing timing only provide an account of contrast within and between languages, they do not give an account of the detail of the phonetic implementation of a particular language, much of which may not have a crucial role to play in contrasting the sounds of a language, but which is nonetheless learned, and part of the phonetic control underlying the production of an utterance.

According to Docherty (1992: 191), one way in which a more detailed account of the timing of voicing could be achieved is by the incorporation of an element of parametric organisation within the descriptive framework in order to provide the level of detail required in describing subtle between- and within-language variability. His model borrows ideas from articulatory phonology (Browman & Goldstein, 1986) in the description of temporal co-ordination of articulatory sequences or 'gestures' in the attempt to incorporate a temporal dimension into phonological representations and to provide a clear representation of the articulatory asynchronies which characterise speech production. Docherty (1992: 86), however, notes that although Browman & Goldstein's (1986) model is one of the few attempts that have been made to overcome temporal resolution problems that are typical of traditional models of phonetic representation, it still lacks a set of language-specific implementation rules between their notion of a gestural score and the coordinative structures needed to realise a given utterance.

Docherty's framework is based on data concerning the timing of voicing in SBE stops and fricatives. The basis of the approach is the incorporation of greater time resolution into descriptions of the timing of voicing by performing finer-grained division of the time base, and specifying the timing of voicing in terms of whether there is any voicing during the medial phase of an obstruent, and the timing pattern occurring at different phases (in obstruent-vowel sequences, vowel-obstruent sequences, and in medial position). The result is a set of possible templates of voicing timing for the different phases and under a range of contexts in a given accent or language. Figure 5.30 shows an example of a template for the timing of voicing in obstruent-vowel sequences:

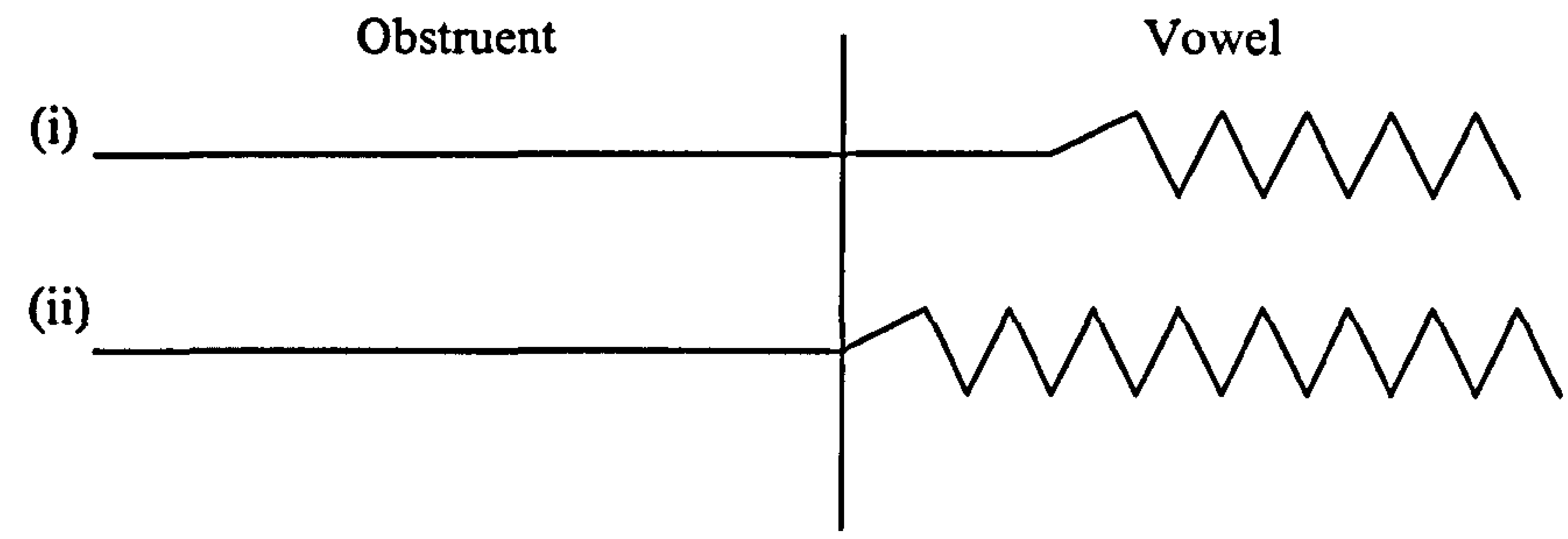


Figure 5.30: An illustration of two of the templates used in Docherty (1992: 193) to represent the timing of voicing in phase 3 (offset) of obstruents in obstruent-vowel sequences in British English.

In type (i), voicing commences some time after the end of the obstruent medial phase. Long delays in voice onset are characteristic of VOICELESS stops, while short delays (of less than 40-50ms) are characteristic of VOICED stops and VOICELESS stops when preceded by /s/ in the same syllable. In type (ii) voicing is present from the end of the obstruent medial phase. This was found in VOICED stops produced by one of the speakers in Docherty's study. Then, a scalar assignment of the continuum follows in order to represent the appropriate amount of detail which would be required for a particular task. A small number of divisions could be used if the aim is restricted to giving a description of the types of contrastive voicing timing patterns used by different languages. On the other hand, the continua could be divided into far smaller portions (such as 10ms units) if the aim is to provide a detailed within- and between- language variability. For instance, the range for the template in phase three described above would be 0-n, with 'n' depending on a number of contextual and language-specific factors.

In order to specify the parameters governing interarticulator coordination, Docherty (1992: 213) calls for a relational rather than absolute specification of temporal aspects of articulatory coordination. This is done by borrowing ideas from Keating's (1990) 'window' model of acceptable configurations located in articulatory space. The position of the window in space and its width are determined on a language-specific basis and navigation through the windows takes place following the most 'cost-effective' way towards achieving auditory goals and their articulatory correlates (Docherty, 1992: 214). In terms of the application of the window model to voicing timing, temporal windows specify the relative timing of laryngeal and supralaryngeal gestures and, depending on the language and the speakers, can be highly constrained in terms of spatial and temporal terms, or relatively underspecified in both dimensions (Figure 5.31).

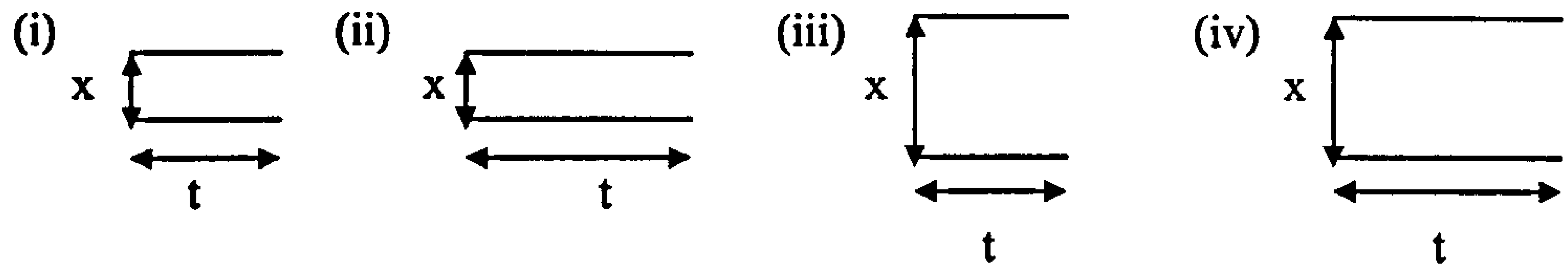


Figure 5.31: Illustration of the means by which temporal windows can interact with configurational windows in order to define sets of acceptable trajectories for a target (x) within the time interval (t) as presented by Docherty (1992: 216).

Spatial and/or temporal windows could be constrained (narrow window) or free (wide window) depending on language specific rules that are governed by phonological and contextual factors. The windows governing voice onset time in utterance initial VOICED and VOICELESS stops in two different languages would be positioned differently reflecting the major differences in the voicing timing of these languages. This would not only cover difference between languages like English and Spanish which use quite different patterns for VOICED and VOICELESS stops, but also fine-grained difference between languages like English and Danish which might seem to use a similar short lag/ long lag contrast but in fact differ in that Danish VOICELESS stops are more aspirated than English ones. The width of a window (i.e. the amount of variability permitted for a particular temporal interval) could also be governed by the importance of a temporal parameter for conveying voicing contrast in a given language, with larger windows possibly reflecting little perceptual weight. Figure 5.32 illustrates how window specification can apply to a representation of voice onset time in English stops (Docherty, 1992: 222).

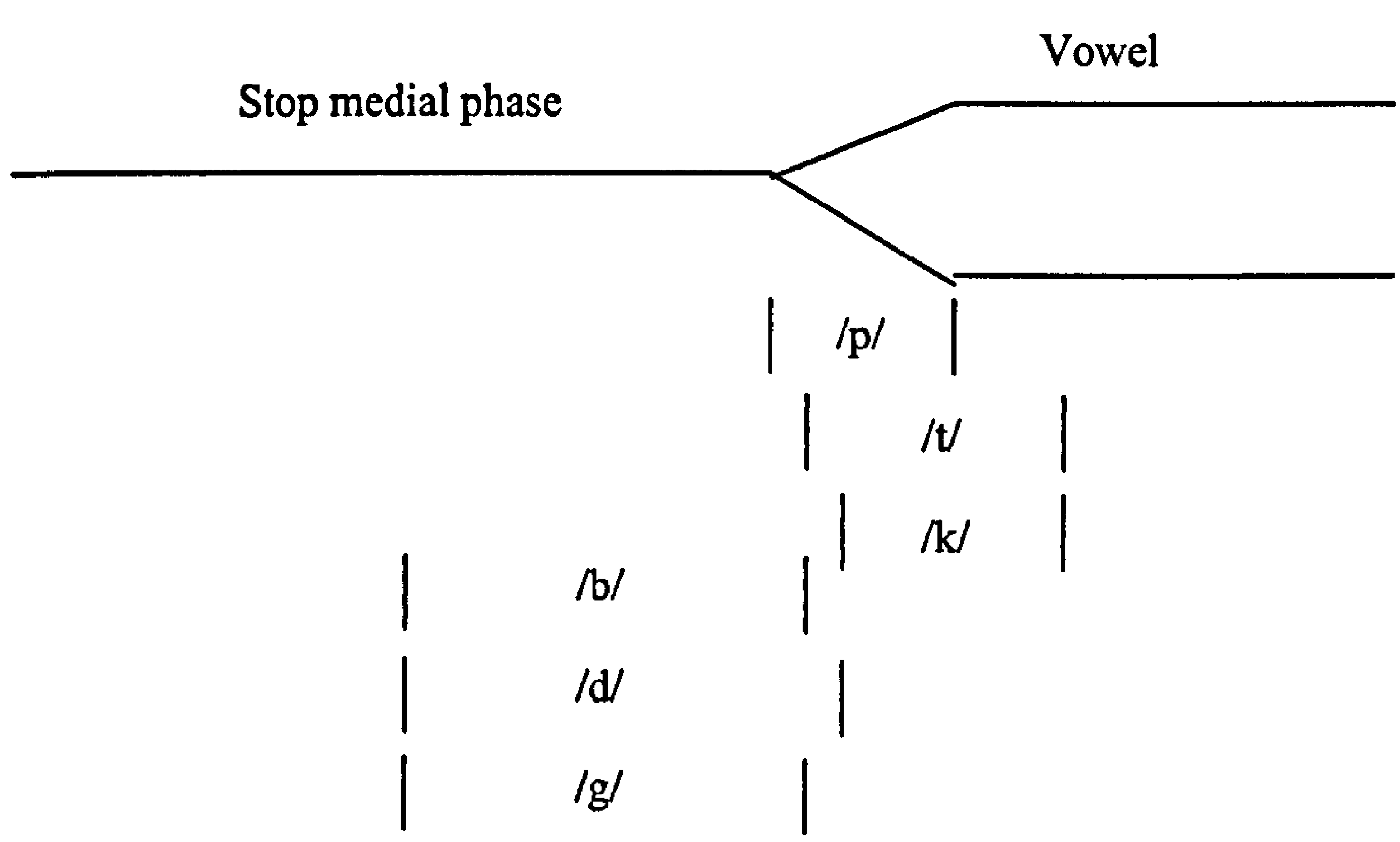


Figure 5.32: A representation of window targets for voice onset time in English stops (from Docherty, 1992: 222).

Although Docherty (1992) maintains that such a proposal is purely descriptive and pertains to the pattern observed in SBE, the model is nevertheless attractive and expandable to other languages and contexts. One advantage is the flexibility of the scalar values to allow a description of timing beyond the 3 abstract categories described by Keating (1984). Another advantage is that the transitional category 'delayed onset of voicing' is not sub-categorised, and therefore aspiration is no longer treated as an all-or-nothing process, but is rather allowed to be a gradient phenomenon (Docherty, 1992: 199).

This observation has recently received support from the large-scale study of VOT in 18 languages that was conducted by Cho & Ladefoged (1999). The authors found that the mean VOT values for unaspirated and aspirated stops in the languages examined occupied the whole positive area of the VOT continuum, ranging from 0 to 154ms. Cho & Ladefoged (1999: 223) note that the line that one draws to separate unaspirated from aspirated stops is arbitrary, and originally suggest four phonetic categories (as opposed to Keating's three categories) to cover unaspirated stops, slightly aspirated stops, aspirated stops, and highly aspirated stops. Following this categorisation, English VOICED stops would fall into the first category while Arabic VOICELESS stops would fall into the second, but the amount of overlap observed in this study means that the categories will always be fuzzy.

However, the authors later mention that there does not seem to be any phonological reason why there might be four groups of categories as suggested, as these do not reflect differences dependent on the number of contrasts in voicing that each language has. Their model (Figure 5.33) therefore falls back to the three traditional categories suggested by Keating (1984), and is an adaptation of Keating's model using some of the notions of articulatory phonology for suggesting language-specific phonetic rules that assign target values for timing between the initiation of articulatory gestures and the initiation of laryngeal gestures (Cho & Ladefoged, 1999: 226). Another adaptation concerns the possibility of specifying more than one target for a given VOT category in the grammar of a given language, which is also possible in Keating's model but is purely decided from context (Keating, 1984: 47).

Although Cho & Ladefoged's introduction of four possible categories was not incorporated into their model, it is relevant for the interpretation of data from this study, which show that there are important phonetic differences between English VOICED and Arabic VOICELESS stops involving divisions that are finer than the boundaries suggested by the three supposedly universal categories, and suggesting that neither of the two broad categories 'short lag' or 'long lag' can adequately describe Lebanese VOICELESS stops.

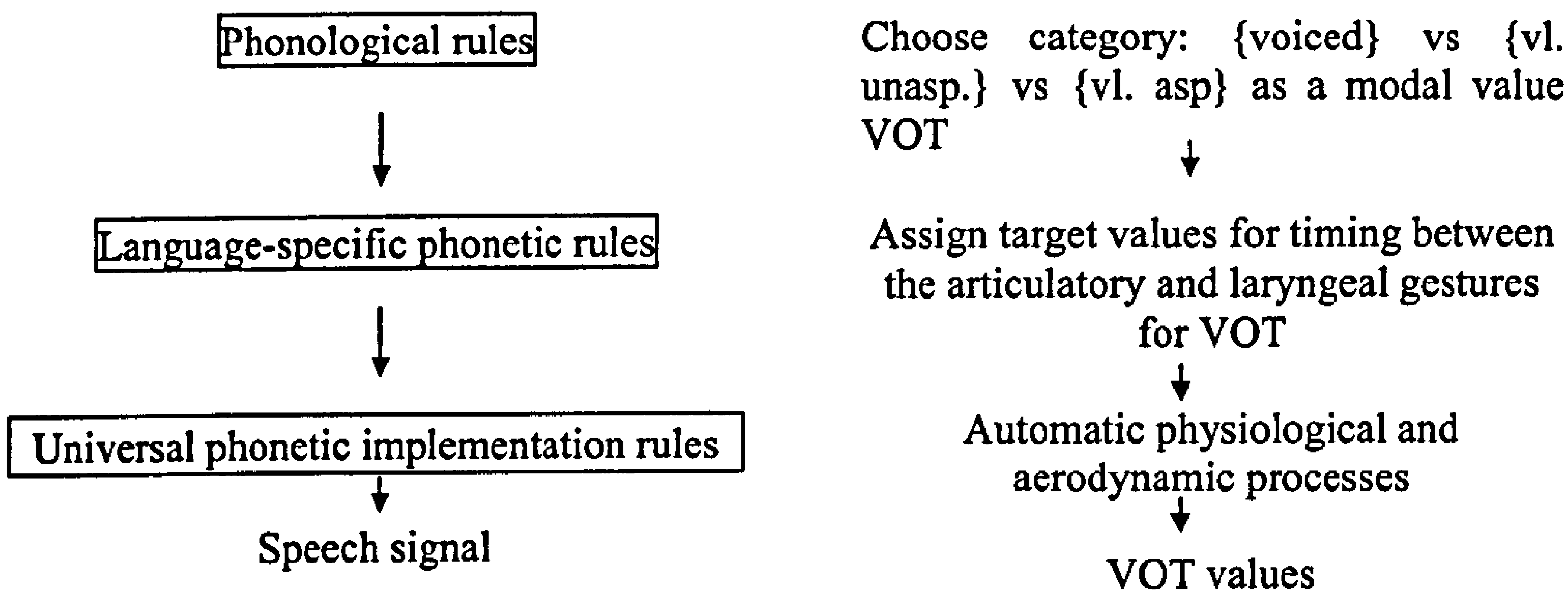


Figure 5.33: Cho & Ladefoged's (1999: 226) representation of multiple processes from phonology to speech signal.

CHAPTER SIX

Discussion and conclusion

6.0 Overview

In Chapter Two, five questions were set for this study in order to investigate how bilingual children learn, store, and use their languages. While each of Chapters Three to Five answered four of these questions in relation to a particular phonological variable, this chapter concentrates on the fifth question, which investigates what the data can tell us about the bilingual's processing of the two languages. Section 6.1 attempts to present the kind of phonological knowledge that the bilinguals from this study might have derived based on the input that was investigated in their environment and its relationship to the bilinguals' output. Sections 6.2 deals with controversial and unresolved issues about language storing in bilinguals and available models of bilingual language processing. While none of the models deals specifically or adequately with the storage and processing of phonological knowledge, an attempt is made to map out the type of phonological information that is stored by the bilinguals using findings from the variables in this study. Discussion of how this knowledge may be used in terms of output and interaction between the two language is then presented in Section 6.3.

6.1 Phonological knowledge

In order to establish a realistic account of the kind of phonological model(s) that is (are) available for the bilingual, one first needs to be aware of the difficulty in doing that in monolingual situations, asexpressed by Vihman (1996: 5):

> "There is, to my knowledge, no evidence that adult speakers of a language share an identical grammar, despite nativist assumptions. On the contrary, individual differences are exhibited in adults as well as children in performance on experimental phonological tasks and in second language learning, retention of spelling patterns and a host of other skills indirectly drawing on phonological knowledge."

This view points to the difficulty in establishing a definable set of target phonological representations for the child to acquire in monolingual situations, let alone bilingual situations. Adult input to the child has been reported as being extremely variable. The sources of this variability may be linguistic (e.g. contextual and coarticulatory changes), or non-linguistic, i.e. related to speaker characteristics (e.g. voice quality, vocal tract length) and social factors such as the speaker's age, gender, geographical background, speaking style, etc. (Docherty et al, 2002; Pisoni, 1997).

Studies reviewed in Chapter One suggest the types of variability described above are not dismissed by any child acquiring the phonology of the ambient language(s). Contrary to what is assumed in many accounts of phonological learning, children do not only acquire the full inventory of adult phonological oppositions, but may also preserve very fine phonetic details and specific characteristics of the speech input (e.g. Docherty & Foulkes, 2000; Foulkes et al, 1999; Local, 1983; Williams & Kerswill, 1999; Roberts & Labov, 1995). This allows them to imitate and reproduce speech patterns heard in their surrounding environment, and therefore provides them with a huge benefit in acquiring the phonology of the local dialect from speakers they are exposed to early in life (Pisoni, 1997: 28; Pisoni & Lively, 1995: 439). Recent research has also shown that the success in acquiring sociolinguistic variability depends on issues such as the complexity of the feature being acquired, the age of the child, and the parents' accent (e.g. Chambers, 2002a; Hewlett et al, 1999; Kerswill, 1996; Roberts & Labov, 1995).

Yet, the tradition in most bilingual research has been to consider the two languages that the bilingual seeks to acquire as consisting of homogeneous sets of well-defined phonological representations (often consisting of a set of abstract phonemes some of which are exclusive to one language while others are 'shared' between the two languages). The child's productions are therefore judged against these targets, and more often than not 'unorthodox' patterns in one language are attributed to influence from the other.

With these issues in mind, I set out to investigate whether English-Arabic bilingual children growing up in Yorkshire acquire language- and accent-specific realizations of /l/, /r/, and word-initial stops, taking into account the sociolinguistic factors that affect their acquisition. These variables were chosen because their acquisition entails not only information that is important for expressing lexical contrast (e.g. VOT), but also aspects of phonetic detail that is essential for the acquisition of socio-phonetic competence (e.g. initial dark [ɫ] in English, /l/ vocalisation, [ɹ] versus [ɾ] realisations of /r/, etc.).

First, it was essential to take account of the input to the children in order to determine the targets that they must be aiming for. The inclusion of the bilinguals' parents along with monolingual children and adults from each language revealed a great deal of variability in the potential input to the children. This in turn offered a substantial contribution to the analysis of the bilinguals' production patterns in the two languages without which some misinterpretations of their linguistic behaviour might have been reached. In Sections 6.1.1-6.1.3, a summary of the targets that are available to the bilinguals for each variable will be presented, along with a discussion of the role of different sources of input and age in shaping the child's developing phonological system(s).

6.1.1 /l/ targets

In English, it was important to make as detailed an assessment as possible of the input that the bilinguals receive before deciding whether they have acquired the appropriate patterns. Although it was expected that Yorkshire /l/'s would be dark-ish in all positions (Wells, 1982), data from the Leeds IViE corpus, the monolingual English friends, and their parents suggested otherwise.

With respect to onset /l/, it emerged that some but not all of the adults from the IViE corpus and the monolingual English parents recorded for this study produced dark [ɫ]'s. For the monolinguals' parents, this outcome is not surprising considering the mixed dialectal background of the speakers, which is representative of the situation in many urban cities where social and geographical mobility result in dialect contact. Contact in turn may lead to what is known as accent levelling, 'a change induced by the reduction of accent features that are socially or locally marked in favour of the adoption of majority features' (Williams & Kerswill, 1999: 149). Levelling has been shown to occur in mobile populations where there is a high level of dialect contact. In such areas individuals regularly find themselves in face-to-face interaction with speakers of other varieties, and in their efforts to accommodate to their interlocutors, tend to avoid local features that are unusual or markedly regional, or which might lead to comprehension difficulties (Trudgill, 1986: 25). But there was also a sign of gender differences with respect to /l/ production in both groups of adults investigated, with the females producing more clear [l]'s than the males. While further investigation of this issue needs to be made, there are three possible interpretations of the children's choices.

First, in cases of dialect contact, Williams & Kerswill (1999: 149) note that first-generation migrants (i.e. the monolinguals' parents in this study) will adapt in minor ways to their new linguistic environment, while second-generation migrants (i.e. the children in this study) produce the phenomenon of linguistic 'focusing', defined as the 'reduction in the amount of linguistic variability in a speech community'. It is not within the scope of this study to offer evidence for focusing, but results for the monolingual English children's /l/ productions in syllable-onset position suggest that the children might have encountered a range of accents and are contributing to a levelled variety, in this case the clear [l]. Only a larger and more representative sample of speakers from the community can help explain the children's choices.

Second, if the children are indeed participating in change that is induced by the dialect contact situation described above, the reason why they might be opting for the clear rather than the dark variant in initial position may be related to articulatory and developmental factors. Clear [l] normally develops earlier than dark [ɫ], and the latter is normally subject to vocalization in final position.

Third, in the three monolingual families that I studied, the mothers are the main caregivers for the children. Since the females were more likely to produce clear [l]'s than the males, the children may still be following the female model, since they are all at the pre-adolescent stage, and teenage group and outside influence has not yet taken over. Chambers (1992) actually suggests that the variants adopted by females have the best chance of being transmitted. The children are, however, showing signs of adoption of vocalized /l/ in coda position, an accent feature that is within the limits of their articulatory abilities and that has been reported as a widespread feature in surrounding regions (e.g. Williams & Kerwsill, 1999). /l/ vocalisation in this study increased rather than decreased in the production of older children, which rules out the possibility that it is simply a developmental feature, especially that vocalisation was also found in the monolingual parents' productions.

What is important for the purpose of this study, though, is that results for the bilinguals (who also mainly produce clear initial [l]'s) turned out to be similar to those of the monolinguals of the same age. Therefore their behaviour should not be interpreted as a failure to produce sociolinguistic aspects of their environment, or as an influence from their parents' L2 productions. The latter option is more obviously ruled out when looking at results for /l/ in coda position, since the bilinguals' parents show signs of language interference by producing a substantial amount of clear [l]'s, which are permissible in Arabic in this context. The bilingual children are therefore exposed to more variability than the monolinguals in this case, since they regularly listen to at least three different realization of coda /l/'s in English (clear, dark, and vocalized) from different sources. Despite this input variability, the bilinguals not only show evidence of having acquired the expected dark variety of /l/ in this context, but also show signs of sensitivity towards sociolinguistic aspects in the host society by producing a considerable amount of /l/ vocalisation. In this respect, the bilinguals may be seen as participants in linguistic change and as contributors to it.

In Arabic, it was important to examine the productions of the monolingual Arabic parents and children to avoid drawing the erroneous conclusion that the bilinguals have not acquired the production of emphatic [lˤ] in Arabic. Emphatic [lˤ] proved to be highly variable in their production due to developmental, contextual, social, and accentual factors. As a result of this variability, it is not surprising that both the monolingual and the bilingual children produced a great number of clear [l]. As for the bilingual brothers (B7 and B10), the fact that they produced more emphatic tokens than the other children and extended emphasis to a context only produced by their parents shows that they have acquired a feature of their parents' accent. It furthermore shows that they have the same ability as monolinguals to acquire emphasis given sufficient input.

Auditory and acoustic investigations of clear [l] in English and Arabic revealed a further subtle differences between the two, in that Arabic clear [l] is actually 'clearer' than English [l] (F2 for clear [l] in English as produced by the monolinguals' mothers and children was significantly lower than F2 for clear [l] as produced by the monolingual Arabic mothers and children in comparable contexts). Though such a finding needs further investigation using a more controlled experiment, it points to the importance of reassessing the concept of cross-linguistic similarity (Strange, 1995). Clear [l] cannot be considered 'the same' in English and Arabic, since the fine acoustic differences that were found between the two in this study may suggest underlyingly different articulatory strategies that are involved in their production. For the bilinguals, this means more variability in the input that they receive, but also means that they have to refine the targets that they are aiming for if they are to produce language-specific clear [l]'s. A preliminary acoustic investigation showed a tendency for the bilinguals to produce higher F2 for clear [l] in Arabic than in English, but the difference between the two distributions was not significant. However, the auditory difference between the two variants is not big and, as Watson (1995) suggests, bilinguals may use different strategies in the production of sounds than the ones used by monolinguals without this being perceptible to native listeners.

6.1.2 /r/ targets

With respect to English /r/'s, it emerged (as expected) that the dominant variant is the alveolar approximant, which was produced by all the monolingual groups examined. Apart from this variant, children also encounter other contextual and dialectal realisations for /r/ in the input that they receive (including feedback from their own productions). For instance, one of the monolinguals' parents frequently produced a labial variant [ʋ]. Unlike vocalised /l/, which seems to have been adopted by the children as an accent feature, [ʋ] seems to be phasing out of the children's production in this study. This shows that developmental features will not persist if there is not enough input to encourage their occurrence. While vocalised /l/ was found in the production of most of the adults from the

IViE corpus and the monolinguals' parents, labial /r/ was more restricted in the adult's production.

Children may also encounter contextual realisations of /r/ such as voiced and voiceless variants (e.g. [pʰɹ̥aɪs] 'price' versus [ɹaɪs] 'rice'), taps (e.g. [ˈvɛɾi] 'very'), fricated /r/'s (e.g. [dʒɹ̝eɪn] 'drain') and so on. However, variability will be always be more pronounced in bilingual children's environment. For instance, along with all the other realisations described above, the bilinguals from this study regularly listen to tap and trill realisations of English /r/ as produced by their parents in contexts where an approximant will be produced by most monolinguals around them.

However, the bilinguals in this study have once again opted for the majority community variants, which again shows signs of the development of appropriate socio-phonetic competence. Moreover, their English accent is non-rhotic, despite the prevalence of post-vocalic /r/'s in their parents' productions. However, this does not necessarily mean that, as Chambers (2002b) suggested, the children have an 'accent filter' which prevents them from noticing features that contribute to a foreign accent in their parents' speech (Section 1.5.7.3). These features do occasionally surface in the children's interactions with monolingual English speakers as found in the study (e.g. tapped /r/'s, clear final /l/'s, etc.). More importantly, the features are heavily used in certain social contexts where the bilinguals consider that acceptable, in this case in code-switches during interactions with the parents. This strongly suggests that the children have learned to produce all varieties and have encoded them in memory, but part of sociolinguistic competence involves deciding which patterns to use in which situations. More discussion of how bilinguals cognitively represent languages will be attempted in the following section.

As for Arabic /r/'s, auditory and acoustic investigation from this study suggest more variability in the adult input than previously suggested in the literature. Apart from the tap and trill variants normally described as realisations of Arabic /r/, it emerged that there is another variant that I called weak tap [ɾ̞] due to its auditory and acoustic qualities. This variant was used more consistently by some speakers than others, and a future investigation is needed in order to find out whether its production correlates with any social, gender or stylistic factors. What is important, though, is that the variability in its production also emerged in the children's output, with a correlation between the adults who produce it the most and their offspring. In the case of the two bilingual brothers, one of them (B7) seems to have adopted the use of the weak tap that was found to be frequent in his father's production (BM7), while the other brother (B10) seems to have adopted the frequent use of the trill, which was found to be frequent in his mother's production (BF7).

6.1.3 VOT targets

With regards to VOT, the overall input that the bilinguals receive consists of word-initial stop realisations with quite fuzzy VOT boundaries between the VOICED and VOICELESS stops as far as English input is concerned, and with conflicting phonetic realisations for a given phonological category when the two languages are considered. In English, VOICELESS stops are produced with aspiration by the monolingual parents and their children, and with short lag to slight aspiration by the bilinguals' parents. VOICED stops, on the other hand, are produced with short lag and occasional prevoicing by the monolingual parents and their children, and predominant prevoicing by the bilinguals' parents. The bilinguals are therefore exposed to stop realisations that span across more or less four phonetic categories (voicing lead, short lag, slight aspiration, and long lag) for English alone, with the short lag category occasionally being ambiguous as to whether it is a cue for VOICED or VOICELESS stops.

In Arabic, the only input that the bilinguals receive is from the parents, and this consists of VOICELESS stops that are realised with short lag to slight aspiration, and VOICED stops that are prevoiced. As was found in Chapter Five, there is a great deal of overlap between VOT values for English VOICED stops and Arabic VOICELESS stops despite the suggestion that they occupy different ranges along the VOT continuum, which means that once again the short lag category is ambiguous as to whether it is a cue for VOICED or VOICELESS stops.

Yet on the whole, the bilinguals do keep the VOT systems of their two languages separate, despite the resulting overlap between phonetic realisations within and across the two languages. What should be kept in mind, though, is that VOT is only one aspect of stop production that the children acquire for each language. The fact that there is overlap between the values for /b d g/ and /p t k/ produced by most subjects in this study even though the stops could still be distinguished through auditory analysis shows that there are other cues for the perception/production of stops that are part of what the child acquires for each language. These are discussed below.

In terms of distinguishing pairs of homorganic stops, a number of studies have shown that there are other equally important acoustic cues that play a role. For instance, Stevens & Klatt (1974) underline the role of formant transitions following voicing onset in the distinction between VOICED and VOICELESS stops in English. However, Lisker (1975) notes that rapid shifts in F1 frequency immediately following voicing onset are helpful but not better 'detectors' than VOT in helping infant discrimination of the two stop categories, and underlines the importance of individual differences with regards to responses to different perceptual cues. Klatt (1975: 695) later suggests the use of five acoustic cues other than VOT for the perception of voicing, including low frequency energy in following vowels, burst loudness, fundamental frequency, segmental duration, and prevoicing. In Arabic, Flege & Port (1981) note that, in addition to a VOT difference, the durations of the stop closure interval of VOICELESS stops are significantly longer than those of their VOICED counterparts and might therefore be part of the cues that listeners use to distinguish between the two sets of stops. What is important to remember, though, is that most studies that have investigated the importance of acoustic cues in the perception of voicing contrast have used synthetic stimuli, and, as Lisker (1975: 1548) points out, we cannot assume that inferences made about perception abilities using speech synthesis must apply to natural speech processing as well.

In terms of production, Scobbie et al (2000) suggest that judging the children's production abilities using cues that are salient in the adult production patterns might lead to missing other important cues that the children are using to achieve voicing contrast. In their investigation of the voicing contrast in the production of monolingual children with phonological disorders (reviewed in Chapter Five), the authors found that one of their subject who was perceived as neutralising the voicing contrast in stops was actually producing a 'covert contrast' by successfully manipulating other acoustic cues in the production of stops, mainly the steepness of spectral tilt immediately following voice onset. Furthermore, studies on bilinguals (also reviewed in Chapter Five) have shown that the experience of a bilingual upbringing can lead to the use of different strategies for the production/perception of stops from those used by monolinguals without these strategies being perceptible to native speakers.

It is therefore important to note that there is more to learning the stop voicing contrast than VOT. Moreover, it should not be assumed that the only purpose behind the child's mastering of VOT patterns in a given language is the acquisition of contrasts. While in English, children acquire three sets of minimal pairs for their stop series (/p b/, /t d/, and /k g/), in Lebanese Arabic, /t d/ and /tˤ dˤ/ are the only true minimal pairs if loan words are discarded, but children learn the appropriate VOT patterns for the whole set, regardless of whether or not the stop has a VOICED/VOICELESS counterpart. Such a difference between English and Arabic reduces the similarity that is drawn between the two sets of stops in the two languages and that is assumed to be a potential challenge for the bilingual child. Even when only /t/ and /d/ are considered, the similarity of the labels typically used in English and Arabic conceals important differences in the production/perception of the two stops. Apart from their different VOT patterns, the two stops are alveolar in English and dental in Arabic, and they are associated with allophonic variations that are context- and accent-specific. While this study has only examined the stops in word-initial position, other contexts might offer further evidence for the disparity between the acquisition of the two stops in each language. Such disparity would make it less likely for the English-Arabic bilingual child to acquire a similar VOT pattern for the two sets of stops in either language. Future studies are needed on the production/perception of English-Arabic bilinguals in order to better understand how they process and master the two different phonetic repertoires.

6.1.4 Parental *versus* societal input

As discussed in Chapter One, the role that parental input in bilingual situations will depend on the type of bilingual family, the social context, and the age of the child. Results from this study have shown that parental influence on the bilinguals' developing phonological knowledge and linguistic choices is more or less restricted to the parents' native language.

Since all of the bilinguals' parents in this study are native Arabic speakers, the bilinguals have mainly acquired their phonological knowledge in English from their environment outside the house. Although the parents do occasionally speak English to the children in the presence of monolinguals, the children do not seem to be influenced by the parents' non-native accent and do not produce any of their parents' accent features in English. This, according to Chambers (2002b), is part of children's sociolinguistic competence. Indeed, the bilinguals in this study not only show signs of having acquired native-like targets for the variables under study, but also produce accent features that are found in the production of monolingual children and adults.

The parents' native language input is, however, crucial and influential, as it is just about the only input that the children receive in Arabic. In this case, the bilinguals do show signs of having acquired their parents' accents and, due to the unique situation of controlled input from a single source, one can easily see how accent features can be transmitted to children by their parents. For instance, despite the difficulty of producing emphatics in Arabic, the two bilingual brothers whose parents produce emphatic glottal stops for otherwise plain glottals have acquired emphasis in this context. Similarly, a correlation was found between the parents' preference for certain /r/ variants (weak taps or trills) and the use of these variants by their offspring.

However, due to the scarce Arabic input that the children receive, certain complex features that require extensive exposure might be acquired late or not at all. For instance, B10 shows no sign of having acquired prevoicing for his Arabic VOICED stops although his monolingual counterpart has. On the other hand, B7's drastic change in his production of VOICED Arabic stops offers many implications for both bilingual and monolingual language acquisition. First, it shows that input plays a crucial role in the acquisition of language-specific phonetic features, and that only a combination of extensive input and early age can lead to a successful acquisition of complex features.

Second, B7's acquisition of voicing lead proves that, when the two conditions of input and age are met, bilinguals will follow similar acquisitional patterns to those of monolinguals, or even 'catch up' with them. In the process of doing so, bilinguals will exhibit developmental patterns that are once more similar to monolingual ones, such as the use of short lag instead of voicing lead (Allen, 1985; Macken & Barton, 1979), or the use of continuants preceding the stops in an attempt to acquire voicing lead. Note that B7's use of a nasal-like sound is not dissimilar to that of Allen's (1985) monolingual French subjects who preceded French VOICED targets with a nasal or a vowel segment that permitted continuous voicing, Macken & Barton's (1980) monolingual Spanish subjects who spirantised their Spanish VOICED stops, or Heselwood & McChrystal's (2000) Panjabi-English subjects who showed both features. In all four cases (Arabic, French, Spanish and Panjabi), the subjects are choosing an articulation that does not involve a complete obstruction of the airstream in order for them to prolong the voicing articulation. In B7's case, there is the added reason that pre-nasalisation of fully voiced stops was also found in his mother's speech. B7's production of implosives can also be considered as another strategy used to maintain transglottal pressure difference, and hence glottal vibration (Heselwood, 1998).

In his discussion of children's acquisition of sociolinguistic features, Kerswill (1996) distinguishes between 'simple' rules, which can be acquired at different childhood stages, and 'complex' rules, which must be acquired at an early age (mainly before the age of six) if they are to be acquired at all. While Kerswill bases his discussion of complexity on several phonologically-, morphologically-, and syntactically-conditioned parameters, a similar measure can be developed for the acquisition of phonetic features. There are features that are simple and that can be easily acquired by the child, and others that need extensive exposure from an early age and that require precise physical coordination because of the demands of aerodynamics and of the timing between glottal and supraglottal articulations.

The notion of 'adequate input', however, is difficult to define or quantify with regards to voicing lead due to a number of factors. First, even when a bilingual child receives regular input in a language that contrasts voicing lead with voicing lag, if the parents or available models do not consistently produce long lead for VOICED stops, then this might affect the child's acquisition of this feature. Such was the case when Deuchar & Clark (1995) found that their Spanish subject's use of short lag rather than voicing lead for VOICED stops was actually not dissimilar from the mother's production. In the case of the subjects in this study, their parents have been found to produce voicing lead consistently, but perhaps the frequency of voicing lead that the bilinguals receive in their daily input is still lower than that for voicing lag. Voicing lag has been described as being more 'acoustically salient' than voicing lead. In a study on the perception of VOT by infants, Aslin, Pisoni, Hennessy, & Perey (1981) found that acoustic information in the negative region of the VOT continuum (i.e. lead) is less salient than in the plus region (lag) for infants aged 6-12 months. Though the findings come from infants in an English-speaking environment, they add to the evidence on the difficulty in acquiring voicing lead.

Second, even if input is available both from parents and society, complex features such as voicing lead might undergo change led by the young bilingual generation. For example, in an investigation of the production of Panjabi stops by young English-Panjabi speakers from Bradford, Heselwood & McChrystal (1999) found that the Panjabi spoken in Bradford is undergoing phonological change in terms of the collapse of the three VOT categories (prevoiced, short lag, long lag) into two by the loss of prevoicing in young speech (age 12-22). The authors note that this phenomenon is not only due to the influence of the dominant majority language (English), but also due to the fact that short lag is less marked, as it involves relaxation of laryngeal tension, a phenomenon reported in linguistic change in other languages, e.g. Proto-Semitic to Arabic (Kenstowicz, 1994: 64; Heselwood, 1996: 32).

6.2 Language storing and processing in bilinguals

Most psycholinguistic studies of language processing in bilinguals have been interested in how bilinguals store and access the lexicons and conceptual representations of their languages, while less attention is given to how bilingual phonological knowledge is represented. In this section some of the models of bilingual language processing that are available in the field are reviewed, and then an attempt is made to draw a preliminary representation of how phonetic/phonological detail from input is stored and learned by the bilingual.

6.2.1 Storing languages and meaning

Studies that have focused on how the bilingual's languages are stored have addressed two central problems. First is the familiar question of whether the lexicon for the two languages is represented separately or together. The second is whether the grammatical system of the second language is stored differently from the first, and how its development interacts with the cognitive structure of the mind.

With regards to the first question, Paradis (1987) mentions four different options to explain storage of the two languages in the brain: (i) the 'Extended System Hypothesis', which suggests that there is no separate storage for each language; (ii) the 'Dual Storage Hypothesis', which assumes that there are separate systems for each language, with a separate set of phonemes, rules and words; (iii) the 'Tripartite System Hypothesis', which assumes that language-specific elements are stored separately and joint elements, such as cognates, together; and (iv) the 'Subset Hypothesis', which assumes the use of a single storage system where links between elements are strengthened through continuous use. This implies that, in general, elements from one language will be more strongly linked to each other than to elements from another language, which results in the formation of subsets which appear to consist of elements from the same language, and which can be retrieved separately. This last hypothesis is attractive due to the fact that it accounts for both bilingual and monolingual storage if one draws an analogy between storing different languages and different varieties of the same language. In both cases, the types of conversations that the speakers find themselves in as part of their daily life will determine the strength of links between elements from different languages/dialects/speaking styles, etc. and allow the speakers to engage in code-switching strategies.

Evidence for any of the above hypotheses, however, has been sought by psycholinguistic studies that are mainly interested in the difference between language storing in late *versus* early bilinguals. The research motive is usually to determine qualitative changes that might ensue from the end of the critical period, but age at which this period is assumed to close reaches nothing like consensus (Bialystok, 2001, 93).

From a neurolinguistic perspective, a major distinction is often made between language acquisition (naturally, or in an informal environment, with the extensive involvement of implicit memory), and language learning (by means of formal methodologies, with learnt and intentionally applied rules, mostly in an institutionalised environment) (Fabbro, 1999: 103). This distinction is important, because a number of researchers have argued that separate cerebral structures are involved, depending on the acquisition processes (emotional systems, cortical and subcortical structures) or learning processes (mainly cerebral cortical areas) (Fabbro, 1999: 108). The general assumption is that an individual can become bilingual at any age; however, at a more advanced age more effort will be necessary to obtain results that are often lower than those reached by children, especially with regards to pronunciation and syntax (Fabbro, 1999: 103). Some of the evidence that has been presented for this assumption will now be reviewed, including the controversies surrounding it.

6.2.1.1 Cerebral organisation of languages

Numerous clinical and experimental studies have been conducted on the cerebral organisation of language in bilingual subjects, but the results have mainly been contradictory. Towards the end of the 1970s, the prevailing idea was that language is organised differently in bilinguals and in monolinguals (Bialystok, 2001: 91). However, subsequent experimental studies have not confirmed this hypothesis. Since mental representations were among the most abstract and impenetrable in cognitive psychology, evidence for the mental organisation conceived by early studies was 'at best inferential and at worst entirely absent' (Bialystok, 2001: 91). The usual data consisted of measures such as reaction time differences to various problems, interferences in performance between different tasks, and behavioural consequences of cortical injury.

More recently, technological advances have offered the possibility of 'observing' brains; it was consequently thought that cognition could be made visible, and that representations could be revealed (Bialystok, 2001: 91). However, recent studies using techniques such as event-related potential (ERP), positron emission tomoraphy (PET) and functional magnetic resonance imaging (fMRI) have not provided any simple answers; instead, they have shown that language organization in the brain is much more complex than previously assumed. For instance, neuroimaging techniques that have attempted to establish the configurations of cortical involvement in representing and processing language have shown that the patterns are less universal than we might have hoped. Although Wernicke's area has traditionally been considered the centre of processing for language comprehension and fluency in monolinguals, recent research (Robertson & Gernsbacher, 1998) using PET scanning data of different comprehension tasks has shown that a significant degree of right hemisphere processing is also engaged for comprehension. Similarly, the role of Broca's area for production turns out to be variable across individuals and difficult to delineate (Bialystok, 2001: 92).

A more moderate conclusion that has recently been offered is that different languages are organised partially in the same areas of the brain and partially in separate and specific areas; hypotheses have been put forward that 'the languages localised in the same cerebral areas at a macroscopic level are represented in distinct neural circuits at a microscopic level' (Fabbro, 1999: 207).

6.2.1.2 Cerebral lateralisation of languages

Another controversial and widely discussed issue in the psycholinguistic studies of bilingualism is the notion that the languages of bilingual speakers are less asymmetrically represented in the cerebral hemispheres than the language of monolingual speakers. The literature on lateralisation for language in bilinguals suggests conflicting positions (see discussions in Bialystok, 2001; Obler, Zatorre, Golloway, & Vaid, 2000; and Paradis, 2000). One is that left hemispheric dominance, which is evident in most monolinguals, applies to bilinguals too. A second proposes weaker left lateralisation for language in bilinguals, while a third maintains that there is differential lateralisation for the two languages. Most of the debate revolves around whether there is greater right-hemisphere participation in the processing of one or both languages in the bilingual than in the monolingual.

Gorlitzer von Mundy (1983 [1959]) was the first to suggest that in bilinguals the mother tongue and the second language have a different lateralisation in the two hemispheres. By studying a bilingual aphasic in 1959, he claimed that the language acquired only at the oral level was represented in both hemispheres, and that the language acquired both in written and oral form was lateralised in the left hemisphere. Albert & Obler (1978) found more cases of bilingual aphasics who had suffered right hemisphere damage (10%) than monolingual aphasics (less than 5%). They concluded that in bilinguals more often than monolinguals linguistic functions are represented in the right hemisphere. Since then many studies have been carried out using the most common techniques of experimental neuropsychology (dichotic listening, tachistoscopic techniques, neuroimaging techniques etc.) in order to find evidence for Albert & Obler's claim, but the results have been controversial. For instance, Rapport, Tan & Whitaker (1983) examined the effect of transient inhibition (Wada test) of both cerebral hemispheres in bilingual subjects on their naming capacity. The four subjects involved were able to name 95% of objects during inhibition in the right hemisphere, whereas the percentage was very low during inhibition of the left hemisphere. Such results contradict the view that language in bilinguals is represented to a greater extent in the right hemisphere.

The right hemisphere is nevertheless known to be crucially involved in the processing of pragmatic aspects of language use (Chantraine, Joanette & Ska, 1998). During the first stages of second language learning, or when L2 is not very well known and is rarely used by the individual, the right hemisphere may be more involved in verbal communication, because beginners tend to compensate for their limited implicit linguistic competence in L2 with pragmatic inferences. A stronger participation of the right hemisphere during verbal communication in L2, however, does not mean that language processes per se are represented in the non-dominant hemisphere (Paradis, 1998). It is rather in this light that one should interpret a series of recent studies using PET and fMRI to study the cerebral representation of language in bilinguals having learned their L2 at school after the age of 7, and for which they had a moderate cerebral representation. Dehaene, Dupoux, Mehler, Cohen & Paulesu (1997), for instance, studied the cerebral representation of both languages during listening to stories in L1 and L2. In the listening condition in L2, the subjects presented on average a greater activation of the right hemisphere then in the listening condition in L1.

6.2.1.3 Methodological problems in neurolinguistic studies

Albert & Obler (1978) identified a number of factors that affected the cerebral dominance of bilingual speakers, possibly accounting for some of the contradictory finding. These included the age and manner of acquisition for the second language, the usage patterns for each language, and language-specific factors. Hence, older learners may represent language differently from younger ones, spoken languages may represented differently from those used only for reading, and some languages may gravitate to the right hemisphere for all speakers. In their view, then, there was not a single configuration for representing languages across two cerebral hemispheres, though their assumption was that the second language always occupied some portion of the right hemisphere.

In more recent publications, Obler, Zatorre, Galloway & Vaid, J. (2000) plead caution in research design and conclusion, and point to the importance of taking into consideration a range of methodological parameters such as language, subject, and stimulus selection, testing procedure, data analysis and theoretical questions about interpreting dichotic and tachistoscopic measures of lateralisation. The authors conclude that findings of differential lateralisation for a set of language stimuli or for a group of bilinguals cannot be understood as 'greater right-hemisphere participation in language processing than normal', i.e. than in monolinguals until all artificial explanations can be ruled out. The complexity of factors involved in the study of language lateralisation in bilinguals must certainly caution us not to assume that any given study can speak for all bilingual individuals, nor for all bilingual populations.

Paradis (2000) sends out the same cautionary notes and goes as far as calling the research on language lateralisation in bilinguals a 'fruitless pursuit'. He wonders why the topic is still as popular as ever after all the contradictory results of the last two decades. Not only can we not generalise to all bilinguals from any given sub-group, we cannot even generalise to any sub-category of bilinguals, no matter how subcategorised by sex, degree of proficiency, age, and manner of acquisition. Paradis (2000: 395) wonders how any of the paradigms used (dichotic listening, tachistoscopic presentation in half visual fields, EEG, etc.) could be a reflection of laterality of language function if so many variables can have an effect on the results. Fabbro (1999: 210) further notes that many methodological issues have not yet been resolved with respect to monolingual subjects, let alone bilinguals.

6.2.1.4 Implications for the nature of linguistic representation in bilinguals

Bialystok (2001: 98) wonders what the implications would be even if we did find that languages are represented differently as a function of being learned first or second. She notes that it is not surprising to find that representations change over time; as competence builds, there is reorganisation of knowledge to accommodate the increasing expertise. The later knowledge will have a different interpretation, a different structure, and a different relation to prior knowledge depending on the state of the existing knowledge. For this reason, it would be surprising if second languages learned at a later stage in life were not represented differently from earlier learned languages in some fundamental way, including perhaps spatial location. Moreover, most of the evidence that the L2 is represented differently is based on the assumption that the first language is localised in the left hemisphere. Although this is generally true, it is not absolutely or universally so (e.g. Satz, 1979, on left- versus right-handers).

Bialystok (2001: 118) offers several suggestions for a better representation of the two languages. First, there is no reason that several organisations cannot coexist, and that these organisations change over time and with development. There is no reason that two representations cannot share certain elements and not others, and that the shared elements cannot change over time. There may well be a range of normal variation that defines how knowledge of language is represented, and across different individuals, that organisation may be quite different. It would not be surprising, for example, if such individual variations reflected differences in experience and language learning history.

These variations might reveal more about the testing methodology than the representation of linguistic knowledge. There is a difference between the structural details of knowledge representation in the brain and the functional use of those representations in processing. The two might not be related: representation that are spatially distinct might be highly interfering during processing and other that are combined may not interact. The resolution might be achieved by examining changes in processing rather than changes in representation as children develop, to which we now turn.

6.2.2 Bilingual language processing

Whether the two language representations are independent or integrated, most current models of language processing assume that both language sources are active when one of them is being used. But if this is the case, then models would have to account for how language performance proceeds fluently in only one of them. There are no theories about the bilingual speaker that aim at a description of the entire language production process (De Bot, 2000: 420). Models have tried to account for how learners can use one language and not the other, invoke the other language when needed, and resist interference from perhaps a stronger language. But a full model that covers the whole process from message generation to articulation is still lacking. In this section, the three most elaborated models are reviewed: Green's (2000) model of inhibitory control, De Bot's (2000) bilingual production model, and Grosjean's (1998) bilingual model of lexical access (BIMOLA).

Green (2000) proposes a general model accounting for verbal expression in bilinguals, which includes data and hypothesis derived from both psycholinguistics and neurolinguistics. The model is based on a modular principle and presupposes the existence of different mutually independent subsystems (e.g. different subsystems for the analysis of words in L1 and L2 and for expression in L1 and L2). These subsystems are in turn formed by an infinite number of independent, yet interacting modules (Figure 6.1).

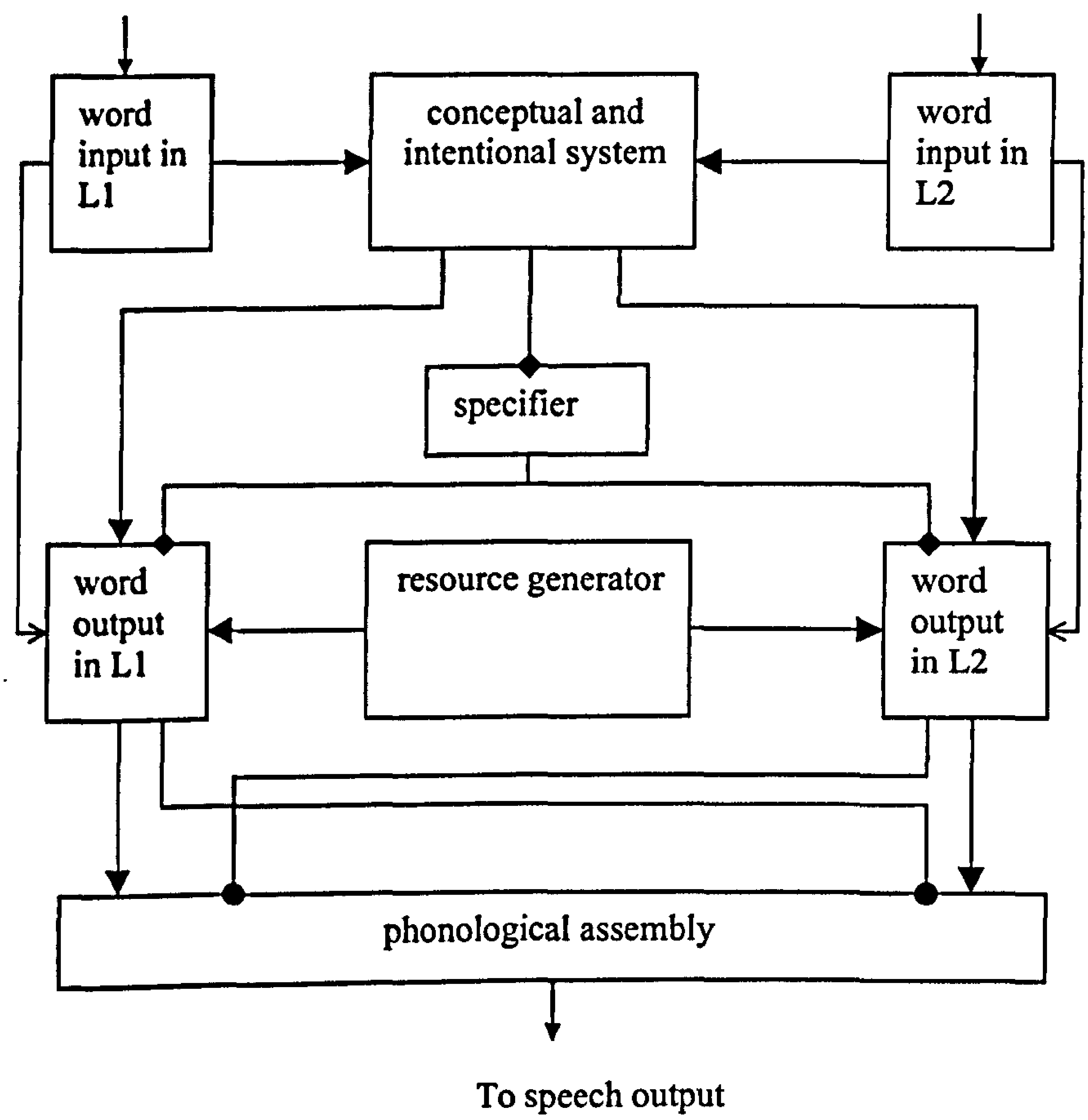


Figure 6.1 Green's (2000) inhibitory model for a bilingual speaker within the control, activation, and resource framework. ——▶ flow of activation; ——◆ control instructions; ——● inhibitory control.

Green's (2000) framework postulates three types of relations between the various subsystems: activation, inhibition, and resource. In terms of activation, each word and each language has a specific activation threshold depending on the frequency of use and on the time elapsed since the last activation. Green (2000: 411) maintains that in most cases both languages remain active, and provides evidence from psycholinguistic experiments such as lexical decision tasks and naming techniques, but also from examples of interference in bilingual production, as well as the ability to code-switch without involving dysfluency. However, the level of activation of a given language can fall if it is not used for a long period, in which case it might reside in long term memory and exert no effect on ongoing processing (Green, 2000: 411).

In the case of the bilinguals in this study, English is more frequently used than Arabic, and it is therefore not surprising that they rarely code-switch to Arabic during the English sessions, but frequently code-switch to English during the Arabic sessions. Furthermore, although not illustrated in his model, Green (2000: 410) acknowledges that the process of word production may be divided into a stage at which the speaker activates words of a certain meaning and a second stage where the actual sound or phonological form of these is retrieved (e.g. Garrett, 1982). This in turn could explain how the bilinguals from this study could access English words during the Arabic sessions due to their frequency, but then apply Arabic phonology to their production.

According to Green (2000) the inhibition component is very important and is also present in monolinguals. For instance, if the word 'apple' is selected, all semantically related words are inhibited (pear, orange, banana), as are all phonologically similar words (e.g. 'dapple'). In bilinguals, if a person wishes to speak one language only, this language is selected and the output from the other language system inhibited. Inhibition is generally automatic and avoids interference between the two languages. Therefore, the activation of the word 'apple' in one language also inhibits the corresponding word in the L2, as well as semantically and phonologically similar words. For selection and suppression to work, the system must be able to identify the relevant outputs. Green (2000: 412) suggests that words possess particular 'tags', where a tag is 'a feature label associated with each individual item'. Some form of tagging may also be used to label vocabulary or structures associated with particular registers or styles of speech within a language. Then, a device called the 'specifier' selects one of the two languages by increasing its activation and suppressing the activation of the other language. Dysfluency in L1 occurs whenever there is an L2 expression of a concept which is more available than one in L1. In the case of code-switching, there is no suppression of L2 and the output can be free to vary according to which words reach threshold first. Switches will obey the syntactic properties of the two languages, although Green (2000: 414) suggests that there is no special device or grammar required to achieve this goal.

The last component is the resource generator, which refers to the amount of energy available for the activation of each of the bilingual's languages. Green (2000: 412) notes that the resources available for the verbal expression in each language are limited, and the task of the resource generator is to replenish the resources at the right rate that is required to control activation and inhibition.

Following Green's (2000) model, part of the children's development in the early years may be in refining this inhibitory control so that they effectively eliminate intrusions from the unwanted language. In the early stages, if both languages are active, two factors may lead to their combination in early speech. The first is that the need to communicate would compel the child to recruit whichever resources are available. The second is that the inhibitory processes required to suppress the non-relevant language might be too fragile to prevent all intrusions for very young children (Bialystok, 2001: 119).

Another model has been suggested by De Bot (1992; 1996), based on the re-adaptation of the speaking model by Levelt (1989). Although Levelt's model was developed to describe monolinguals, De Bot notes that many aspects of speaking are the same for monolingual and bilingual speakers, and therefore a single model to describe both types of speaker is to be preferred over two separate models. De Bot's model hypothesises the existence of (i) three subsystems for language production (a conceptualiser, a formulator, and an articulator), (ii) a subsystem for comprehension, and (iii) the lexicon, a subsystem that is involved in both production and comprehension (Figure 6.2).

The conceptualiser contains all the information that can be expressed by means of language but that is not linguistic itself (preverbal messages). Levelt notes that the conceptualiser contains a 'discourse model', or a list of conditions for the speech which is to be generated, including the use of 'registers', which Levelt (1989: 368) defines as 'varieties which may have characteristic syntactic, lexical and phonological properties. De Bot (2000: 427) notes that adopting 'registers' could be generalised to 'varieties' and 'languages' and that, therefore, there is no difference between the different registers used by a monolingual speaker and the languages spoken by a multilingual speaker. In bilingual subjects, the conceptualiser is therefore responsible for conventions in conversation, which are language-specific, and for choosing which language should be used in a given utterance (De Bot, 2000: 427).

The formulator converts the preverbal message into a speech plan (phonetic plan), by selecting the right words or lexical units and applying grammatical and phonological rules. Lexical units consist of two parts: the lemma and the morpho-phonological form or lexeme. In the lemma, the lexical entry's meaning and syntax are represented, while morphological and phonological properties are represented in the lexeme. In production, lexical items are activated by matching the meaning part of the lemma with the semantic information in the preverbal message. While the surface structure is being formed, the morpho-phonological information belonging to the lemma is activated and encoded. The phonological encoding provides the input for the articulator in the form of a phonetic plan (De Bot, 2000: 423).

In the bilingual, the division of the lexical units into two parts probably provides an explanation for lexical items that are borrowed from one language and adapted to the phonology of the base language. But the question is whether bilingual individuals have (i) a formulator and a separate lexicon for each one of the known languages or (ii) a unique large system that stores all data concerning the different languages (De Bot, 2000: 428). De Bot notes that factors such as the linguistic distance between the two languages and the level of proficiency involved will affect the organisation of the formulator and the lexicon. Fabbro (1999: 214) adds that age and method of acquisition as well as the use of a language all play a role. But as mentioned in Section 6.2.1, there is no conclusive evidence regarding the relationship between issues like age and proficiency and language storage and representation in bilinguals.

The articulator converts the speech plan into actual speech. Levelt assumes that the syllables are the basic units of articulatory execution. In this view, phonetic plans for words consist of a number of 'syllable programmes' so that the speaker has an inventory of syllables that need not be generated from scratch every time a word is produced. The phonetic plan therefore consists of a string of syllable programmes (De Bot, 2000: 435). For the bilingual speaker, the situation may depend on the proficiency attained in the two languages, as syllable programmes are assumed to be automised, and the level of automaticity is likely to correlate with the level of proficiency. Also in this model, when the number of syllables to be stored may become large, it is assumed that syllable-programmes that are the same for two languages will not be stored twice, while language-specific ones will be uniquely represented.

Problems with this approach can easily be spotted, as there are no simple grounds for establishing similarity in the two languages. The phonological encoding module also contains a prosody generator, which constructs a temporal structure and a pitch contour for the utterance (Levelt, 1989: 398). For bilingual speakers, De Bot (2000: 437) suggests that there is one articulator which has an extensive set of sounds and pitch patterns from both languages to work with.

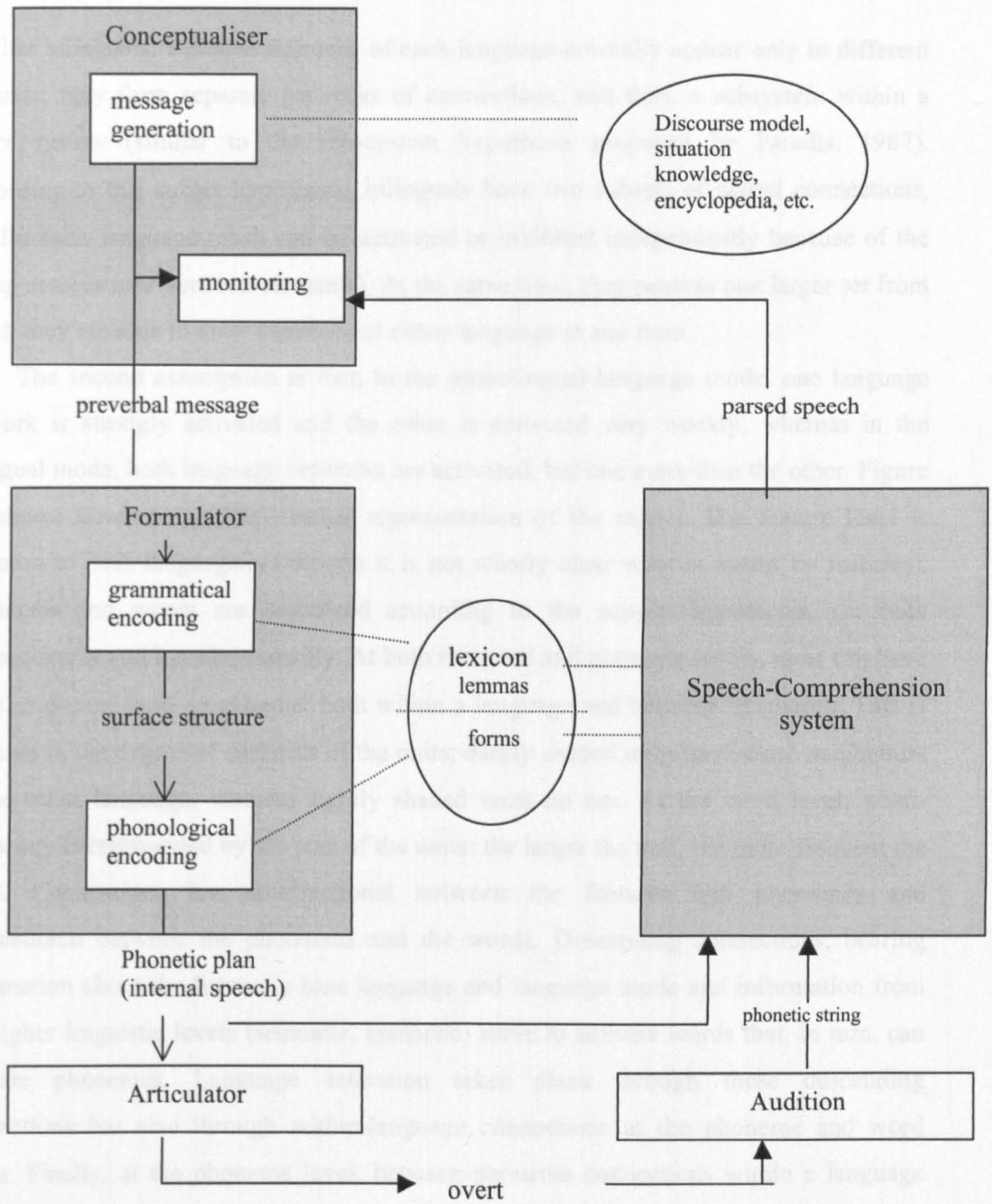


Figure 6.2 Levelt's (1989) speech production model.

Grosjean (2000) proposed an interactive activation model of word recognition in bilinguals which has since been named BIMOLA (Bilingual Model of Lexical Access). It was strongly inspired by the TRACE model proposed by McClelland & Elman (1986) and is controlled by two basic assumptions. First, Grosjean (2000: 466) assumes that bilinguals have two language networks (features, phonemes, words, etc.) that are both independent and interconnected. They are independent in the sense that they allow a bilingual to speak just one language, but they are also interconnected in that the monolingual speech of bilinguals often shows the active interference of the other language and in that bilinguals can code-switch and borrow quite readily when they speak to other bilinguals. Because elements of each language normally appear only in different contexts, they form separate networks of connections, and thus, a subsystem within a larger system (similar to the sub-system hypothesis proposed by Paradis, 1987). According to this subset hypothesis, bilinguals have two subsets of neural connections, one for each language (each can be activated or inhibited independently because of the strong associations between elements). At the same time, they possess one larger set from which they are able to draw elements of either language at any time.

The second assumption is that, in the monolingual language mode, one language network is strongly activated and the other is activated very weakly, whereas in the bilingual mode, both language networks are activated, but one more than the other. Figure 6.3 shows Grosjean's (2000) visual representation of the model. The feature level is common to both languages (although it is not wholly clear what is meant by features). Phonemes and words are organized according to the sub-set hypothesis, i.e. both independently and interdependently. At both the word and phoneme levels, units can have close or distant form neighbours, both within a language and between languages. This is depicted in the degree of darkness of the units; darkly shaded units have close neighbours in the other language, whereas lightly shaded units do not. At the word level, word-frequency is represented by the size of the units: the larger the unit, the more frequent the word. Connections are unidirectional between the features and phonemes and bidirectional between the phonemes and the words. Descending connections, bearing information about the listener's base language and language mode and information from the higher linguistic levels (semantic, syntactic) serve to activate words that, in turn, can activate phonemes. Language activation takes place through these descending connections but also through within-language connections at the phoneme and word levels. Finally, at the phoneme level, between-phoneme connections within a language can allow for phonotactic activation.

Grosjean (2000: 467) claims that, in the bilingual mode, the activation unit in one network and of its counterpart in the other depends on their degree of similarity. Within this view, if English /b/ is activated at the phonemic level, then French /b/ is activated as the consonants are considered similar. On the other hand, the activation of English word-initial /p/ will lead to a much lower level of activation of French word-initial /p/, as the two consonants are considered quite different. When English /r/ is activated, its French counterpart should receive very little activation.

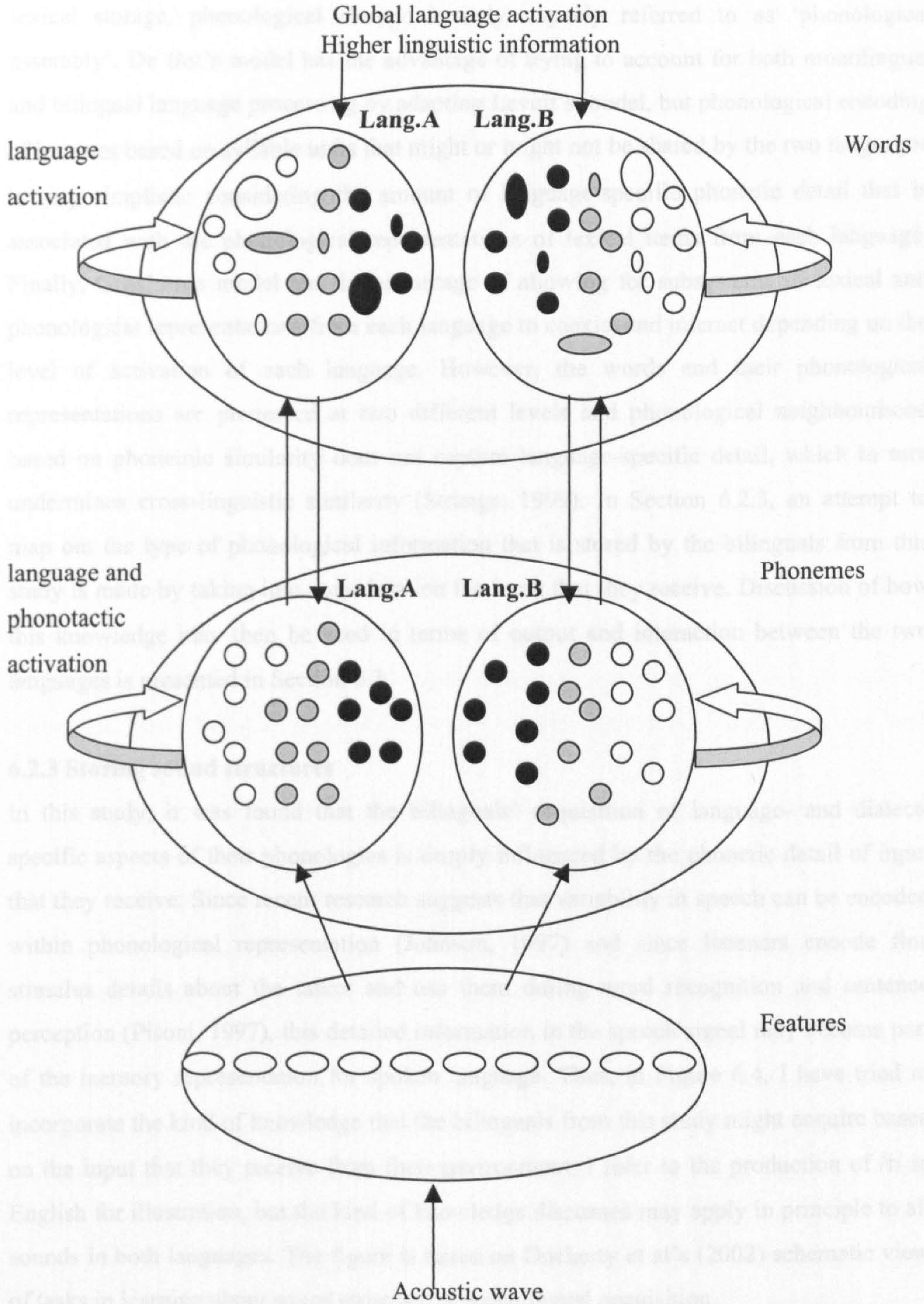


Figure 6.3: Grosjean's (2000) Visual representation of the BIMOLA model of lexical access in bilinguals.

While the three models reviewed in this section tackle different important issues related to bilingual language processing and lexical access, the way phonological knowledge is represented in bilinguals remains simplistic and removed from the type of input that the bilinguals receive. As Green's (2000) model is mainly concerned with lexical storage, phonological storage is only vaguely referred to as 'phonological assembly'. De Bot's model has the advantage of trying to account for both monolingual and bilingual language processing by adapting Levelt's model, but phonological encoding of lexemes based on syllable units that might or might not be shared by the two languages is very simplistic considering the amount of language-specific phonetic detail that is associated with the phonological representations of lexical items from each language. Finally, Grosjean's model has the advantage of allowing for subsystems of lexical and phonological representations from each language to coexist and interact depending on the level of activation of each language. However, the words and their phonological representations are processed at two different levels and phonological neighbourhood based on phonemic similarity does not capture language-specific detail, which in turn undermines cross-linguistic similarity (Strange, 1999). In Section 6.2.3, an attempt to map out the type of phonological information that is stored by the bilinguals from this study is made by taking into consideration the input that they receive. Discussion of how this knowledge may then be used in terms of output and interaction between the two languages is presented in Section 6.3.

6.2.3 Storing sound structures

In this study, it was found that the bilinguals' acquisition of language- and dialect-specific aspects of their phonologies is deeply influenced by the phonetic detail of input that they receive. Since recent research suggests that variability in speech can be encoded within phonological representation (Johnson, 1997) and since listeners encode fine stimulus details about the talker and use them during word recognition and sentence perception (Pisoni, 1997), this detailed information in the speech signal may become part of the memory representation for spoken language. Thus, in Figure 6.4, I have tried to incorporate the kind of knowledge that the bilinguals from this study might acquire based on the input that they receive from their environment. I refer to the production of /r/ in English for illustration, but the kind of knowledge discussed may apply in principle to all sounds in both languages. The figure is based on Docherty et al's (2002) schematic view of tasks in learning about sound structure in monolingual acquisition.

From type A input, a child learning English deduces information about lexical contrast, including the semantic and phonological distinction between words (e.g. 'ran' *versus* 'ban'), and phonotactic rules specific to the language. It is this sort of information that is usually considered crucial to the development of the child's phonological system.

The bilingual child in this case, may be considered to acquire a phonological contrast between English [ɹan] 'ran' and Arabic [r̥ăn:] 'he rang'. However, as can be seen from the example provided, apart from the difference in /r/ production in each language, the following vowel and nasal also vary in their production between English and Arabic. One therefore needs to keep in mind that there might be subtle acoustic and articulatory differences that are not represented by IPA symbols but that are part of what the language-specific detail that the bilingual acquires (e.g. clear /l/'s in English and Arabic are presented using the same symbol but may vary articulatorily and acoustically). The notion of 'analogous phonemes' in studies of cross-linguistic similarities therefore faces the problem that different languages have dissimilar sound systems in many respects (Scobbie, 2002). Moreover, the different sound inventories, prosodic systems, phonotactics, morpho-phonological patterning of sounds in two languages make comparisons of 'similar sounds' very difficult and makes it less likely that the bilingual will substitute 'similar' segments between the two languages.

Type B input illustrates what might traditionally be called sociolinguistic knowledge reflecting variation in pronunciation linked to age, sex, etc. This is clearly knowledge that has to be learned, but is usually considered outside the scope of phonology itself. The bilinguals in this study will learn to associate tapped /r/ productions with input from their parents, and approximants with input from monolinguals in their environment.

Type C input deals with both kinds of knowledge, as bilinguals will experience input forms which simultaneously encode contrastive and sociolinguistic information. In the case of /r/ production, examples such as [bɪə] and [biːɾ] will be present in the input. On the one hand, these forms provide information about potential phonotactic distribution of /r/. On the other, the alternative forms also have clear sociolinguistic associations since the latter will only be produced by the children's parents. The bilinguals may therefore learn multiple representations of sounds and associate them with particular speakers, languages, styles, situations, etc. As Docherty et al (2002) point out, children may look for sound-meaning associations of all sorts within the ambient sound patterns without excessive privilege being assigned to lexical meaning. In light of this, Docherty et al (2002) suggest that children might start off with a single assimilated store of knowledge, containing information about phonological contrast and sociolinguistic information that is encoded phonetically (hence the dotted ellipses showing overlap in Figure 6.4). Subsequently, the two types of information may gradually become separated (although Docherty et al also entertain the possibility that some degree of overlap remains permanently).

For bilinguals, we expect that the extent of sociolinguistic information is even greater than that for monolinguals. For example, they have available to them a variety of phonological models for /l/, /r/, and VOT in each language due to the fact that they are exposed to language input from home and from society. This input varies between standard, non-standard, and non-native varieties from different speakers and different contexts, and there is an uneven amount from each variety depending on the speakers that the bilingual is exposed to the most. Therefore, the phonological representations that the bilinguals will develop will be different from either monolingual model (which we know is variable in itself), and from other bilingual models.

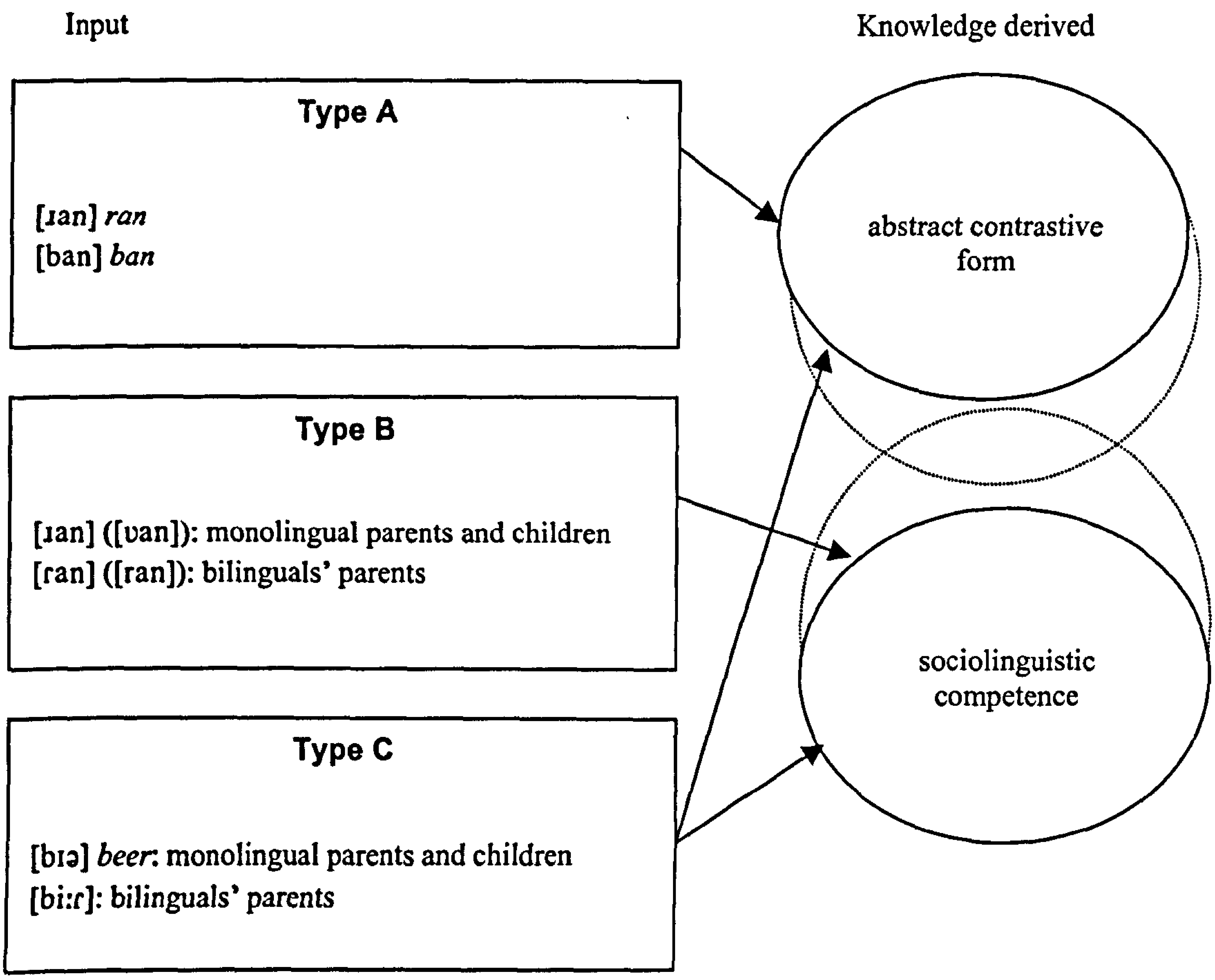


Figure 6.4: Schematic view of tasks in learning about sound structure by bilinguals. Adapted from Docherty et al (2002).

What this section has served to establish is that input contains overlapping sources of information; some information is about lexical contrast, while other information is sociolinguistically relevant. It may or may not be the case that sociolinguistic information is learned simultaneously, as Docherty et al, 2002 suggest, but both types of information clearly must be learned. If we take the broad view of what 'phonological knowledge' is, i.e. knowledge about the production and perception of sounds, then sociolinguistic information becomes all the more relevant for bilinguals, as they learn to associate the perception and production of sounds with different sources of input and different sociolinguistic contexts. There is therefore little room to assume overlap between the bilinguals' phonological systems. This becomes more obvious when one considers that the sounds that the bilinguals learn in their two languages differ not only in their detailed phonetic features, but also in the phonological and phonotactic rules that govern their production, and the social overtones that are associated with their use.

6.3 Language use by bilinguals

In Chapter One, a discussion of the concept of language mode was presented and it was suggested that the state of activation of the bilingual's languages operates along a continuum ranging from monolingual to bilingual depending on who the bilingual is speaking or listening to, the situation, the topic, and so on, and that each language mode will have an impact on the bilingual's production (Grosjean, 1998). This study supports this hypothesis from a phonological point of view, and points to the fact that some language modes that operate in bilinguals are the result of their linguistic background and therefore differ from one child to the other. Figure 6.5 shows Grosjean's representation of the language mode continuum again, so that we can discuss the relationship between the different positions along the continuum and the types of interaction that the bilinguals from this study find themselves in.

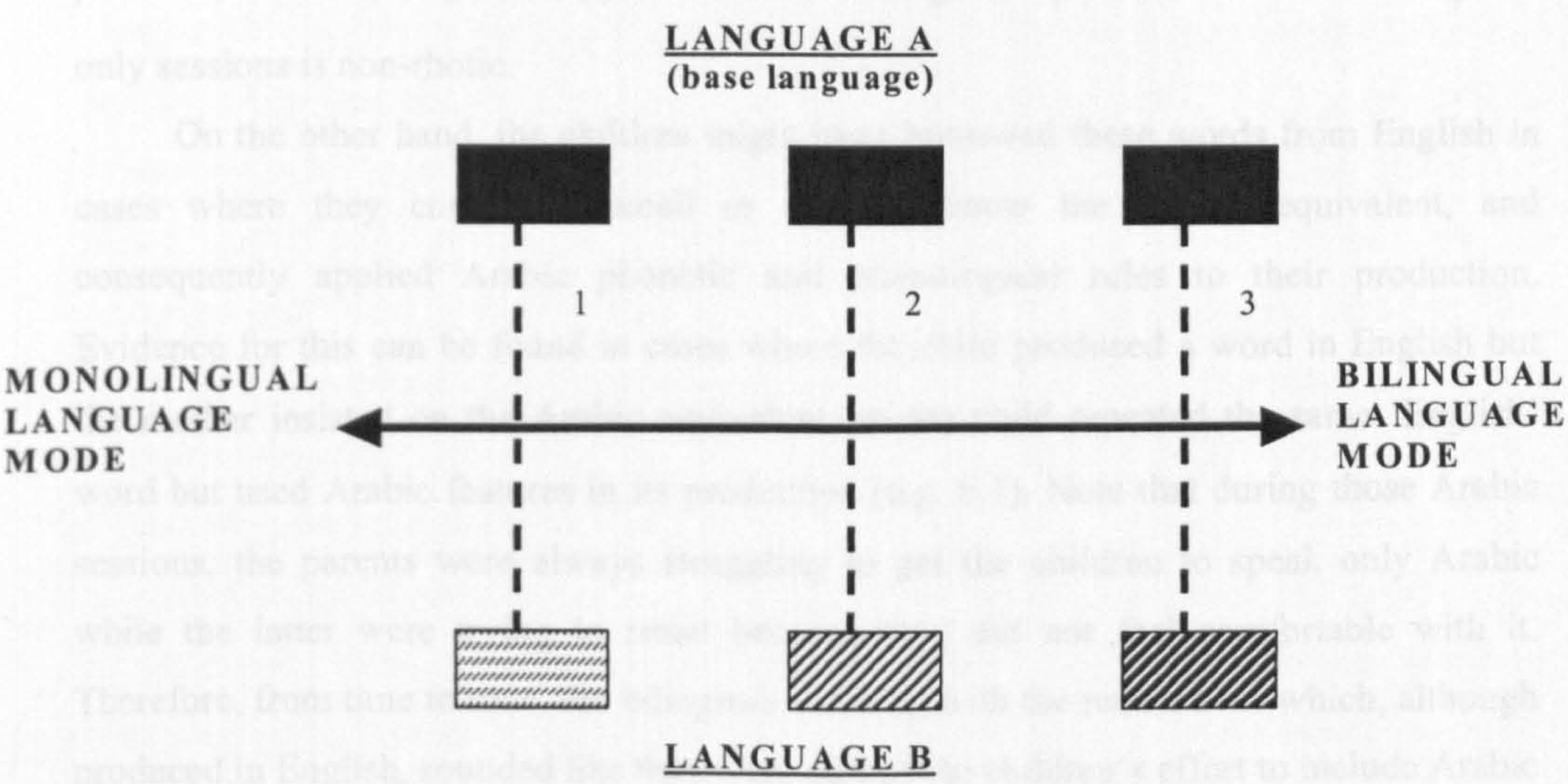


Figure 6.5: Visual representation of the language mode continuum as presented in Grosjean (2000: 3).

The base language that the bilinguals from this study regularly find themselves in is English, since they live in the UK and frequently interact with monolingual speakers of English. The level of activation of the subjects' languages can therefore be considered to be frequently close to position one on the graph. Note that even when two of the bilinguals were taped in free-play sessions together, they mainly spoke English and rarely code-switched to Arabic. The three subjects are English-dominant, and, as a result of that, very little influence from Arabic was found in their English production, regardless of whether the sessions were conducted with a bilingual (myself) or a monolingual.

Arabic, on the other hand, is the subjects' weaker language, and results from this study suggest that the bilinguals frequently find themselves in position three along the continuum, as their English remains strongly activated when Arabic is the base language. Evidence was found from the frequent code-switching that took place during the Arabic sessions. While the English /l/, /r/, and VOT variants that the children produced during the English-only sessions differed from those produced during the Arabic sessions, the English tokens that were produced in the Arabic sessions contained a mixture of features that belong to both languages. There are different ways to explain this phenomenon. On the one hand, the children might not be aware that these are English words, as they might simply have learned their parents' pronunciations of these words. Evidence for this stems from the fact that the five-year-old, who has limited reading skills, produced words with post-vocalic /r/'s like [ˈmaɾbəz̧] for 'marbles' though her pronunciation in the English-only sessions is non-rhotic.

On the other hand, the children might have borrowed these words from English in cases where they could not recall or did not know the Arabic equivalent, and consequently applied Arabic phonetic and phonological rules to their production. Evidence for this can be found in cases where the child produced a word in English but the mother insisted on the Arabic equivalent, so the child repeated the same 'English' word but used Arabic features in its production (e.g. 6.1). Note that during those Arabic sessions, the parents were always struggling to get the children to speak only Arabic while the latter were trying to resist because they did not feel comfortable with it. Therefore, from time to time, the bilinguals came up with the realisations which, although produced in English, sounded like they were part of the children's effort to include Arabic features in order to please the mothers while still using English. However, not all the children's English productions in Arabic were borrowings, as some of these productions displayed a change not only at the lexical but also at the phonetic level; this type of switch is reported by Grosjean (2000: 454) as being possibly due to the flexibility of the production mechanism.

(6.1)	Mother (pointing at a castle):	[ʃu haɪda]? *What that (masc.)?* 'What is that?'
	Child:	[ˈkʰasɬ̩] <u>CASTLE</u>
	Mother:	[laʔ bɪlˈʕaɾabe] *No in-Arabic* 'No, (say it) in Arabic'
	Child:	[ˈka̹səl]

A third possible explanation is that the children were accommodating to the mothers' productions in English. Evidence for this option stems from observing the children explaining English utterances to their mothers or repeating an English utterance slowly after detecting misunderstanding on the part of the parents. Evidence also comes from the fact that some of the productions that the bilinguals made included features that are not only common in Arabic, but are also particular to the idiosyncratic L2 features that the parents produced and that the children would most likely have heard from them (e.g. [a] for schwa in 'waiter' [ˈweɪtaɾ], geminate /l/ for 'umbrella' [ʌmˈbrɛlla], etc.). The fact that the bilinguals produced these realisations only during the Arabic sessions and not the English ones suggests that they may have imitated the parents' English productions. B7 and B10's parents also note that the two brothers often use 'more complicated English' when they are playing together than when they are speaking to their parents. Accommodation is a sign of communicative competence that has often been mentioned as being part of the behaviour of the bilingual (Fantini, 1985: 116; Hamers & Blanc, 2000: 253; Hoffmann, 1991: 180). Bilinguals are known to be able to 'accommodate' their speech according to the needs of their interlocutors. This can take place by either choosing the language that suits the participant or, within the chosen language, adapting the speech to the level of the listener (e.g. speaking slowly, emphasising the pronunciation of words).

Regardless of the reasons discussed above for the code-switches and borrowings, the main point to be made in this study about the phonetic patterns that were found in these code-switches and borrowings is that they cannot simply be considered a result of interference between the two languages of the bilinguals, as they only apply to the English produced in the Arabic sessions. It is important to view such patterns as the product of strategies employed by children to enhance communication in their weaker language (Grosjean, 1982: 191). Due to the adoption of principles from second language acquisition by many bilingual researchers, interpretations of the bilinguals' developing languages have often been given terms such as 'interference' or 'transfer' to refer to the influence that one language might have on the other. From this perspective, whenever the bilinguals' production in either of their languages is compared, any 'unexpected' pattern that is not 'similar' to monolingual patterns in each language is interpreted in terms of possible influence from the other language. If the language mode is accounted for, then only deviations from the norms that cannot be controlled by the bilingual and that take place during conversations with monolingual speakers could be considered interferences. For the subjects in this study, interferences were in the form of clear final [l]'s and taps in English that were produced during interactions with the monolingual English children. These were small in number and occasionally occurred in the monolingual children's productions. As Watson (1995) observes, it is difficult to be sure that a bilingual is doing something that a monolingual would never do, as monolingual norms are themselves constructs that conceal potentially wide variation.

Grosjean (1982: 293) notes that even if a bilingual has the language competence of a monolingual in both languages, he or she will rarely be able to keep the two languages completely separate when talking to a monolingual; from time to time, they will influence one another, even if only momentarily. Factors such as fatigue, stress, topic of conversation, situation, and the interlocutor will affect the frequency of 'deviations'. When the bilingual has achieved a stable level of fluency, breakdowns are much less frequent. Indeed, despite the fact that in this study bilinguals' systems are still developing, the Arabic influence on their English production does not cause any break down in communication and, in fact, often goes unnoticed. Recall from Chapter Two that impressionistic judgements from the bilinguals' teachers and a group of native English listeners revealed that most the listeners were confident the children had a native accent, while only a small number of listeners spotted certain non-native features in their speech. Watson (1995: 38) notes that it is possible for bilinguals to use different production routines from monolinguals in their two languages without being perceptible to other native speakers.

Figure 6.6 shows a schematic representation of the kind of overlapping systems of language use that the bilinguals from this study possess, based on the limited number of interactions with friends and family that were observed. Of course, in reality, the number of those systems will be much bigger, to accommodate for other types of interactions that the bilinguals engage in every day. As can be seen in Figure 6.6, the English spoken by the bilinguals during interactions with monolingual English speakers will not be very different from that used with other bilinguals (thus the nearly overlapping ellipses), due to the fact that it is their dominant language and that they actively choose to speak it with bilinguals and monolinguals alike. The only reason the two systems were not drawn in such a way as to be completely overlapping is that, during the free-play sessions between the two bilingual brothers, there were sporadic code-switches to Arabic. Next, the Arabic spoken by the bilinguals during the Arabic sessions (Figure 6.6, bottom ellipsis) shows very little overlap with the two types of English interactions just described, as the bilinguals kept the patterns of the two languages quite separate. The English code-switches and borrowings (Figure 6.6, middle ellipsis), on the other hand, show a degree of overlap with both the English and the Arabic patterns, depending on whether the bilinguals adapted the English production to Arabic phonology or not.

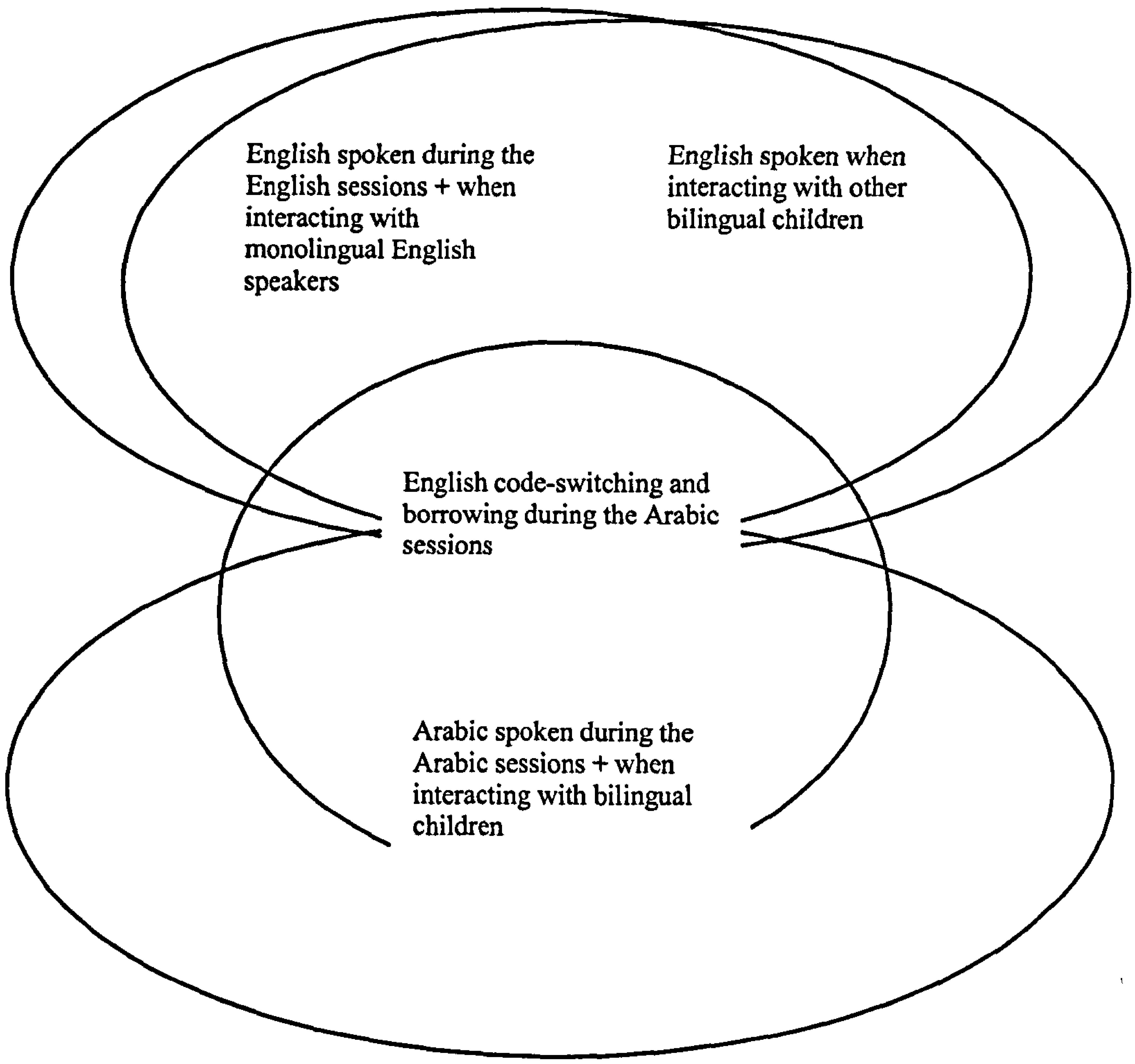


Figure 6.6: Schematic representation of the overlapping patterns of interactions that the bilinguals engage in.

6.4 Conclusion

The results obtained from this study offer important observations related to methodological issues in the study of bilingual phonological acquisition specifically and phonological acquisition in general. First, any examination of bilingual speech needs to take account of the difficulty in specifying the phonological targets that are available to the bilingual for each language. Bilinguals are exposed to input that normally ranges between standard, non-standard, and non-native varieties; these varieties consist of overlapping phonological systems that create fuzzy boundaries for a given phonological target. Even in monolingual situations, a speech community consists of multiple overlapping sound systems, reflecting non-linguistic factors such as gender, age and others which influence the social interactions of speakers (Docherty et al, 2002; Scobbie, 2002). An account of the knowledge that is acquired by the child that is based on multiple-trace models may allow the encoding of such variability within the acquired phonological representations. Each speaker's knowledge of their language will therefore consist of a personal system compiled from their unique experience of the output from other systems that are more or less similar to each other. In bilingual situations, there will be two sets of systems for the child to choose from. Following these considerations, the mental representation of two languages for a bilingual is clearly different from that of a monolingual but certainly not the simple combination that would result from compiling two systems into a place normally assumed as being occupied by one. The languages of bilingual children need not be, nor are they likely, to develop entirely autonomously or interdependently. Certain aspects might develop interdependently, while the rest develops autonomously (Genesee, 2001: 159).

Second, variability is also recognised as one of the most obvious characteristics of children's speech, so this issue needs to be taken into consideration when interpreting bilingual children's production. Developmental factors are exerted on all phonological representations as children enlarge their knowledge of language and its structure, expand their resources for using language and communicating effectively, and mature in their social interactions. Similarly, bilingual children draw on their multiple representations that pertain to knowing and using language as they continue to develop more complete representations for each language they are learning (Bialystok, 2001: 120). Like monolingual children, they make do with whatever linguistic resources they have available to express themselves, the only difference being that, unlike monolingual children who are limited to the resources of one language, bilingual children can draw on two (Genesee et al, 1995: 629). Linguistic development is a continuous process, sensitive to the context and the sociolinguistic circumstances around the child. Ultimately, the order and rate of acquisition of one or more languages lies on environmental, social, and psychological factors, and depends on the amount and quality of the input the child receives from the environment with respect to the linguistic forms. Given sufficient exposure to two languages, bilingual children can reach the proficiency level in each of their languages as monolinguals in the long run (White & Genesee, 1996).

Third, children acquiring the phonology of their language(s) do not only gain knowledge of lexically-contrastive phonological features, but also incorporate sociophonetically relevant aspects of linguistic competence (Docherty et al, 1997; 2002). As Schieffelin (1990: 17) points out, the processing of linguistic knowledge goes hand in hand with the processing of social knowledge. Bilingual children learn how to become members of their community and to communicate effectively with different interlocutors from different language and social backgrounds. Based on findings from recent studies of sociolinguistic variation (e.g. Docherty et al, 1997; 2002; Scobbie, 2002; Thomas, 2000), the results from this study confirm the productive outcome of linking experimental phonetics and sociolinguistics for a more refined description of language- and dialect-specific phonological patterns. The difficulty that was involved in conducting instrumental analysis using 'home' as opposed to 'laboratory' speech was outweighed by the benefit of more natural speech and therefore a more realistic representation of the patterns that are normally produced by the speakers.

Fourth, the context in which bilinguals produce their languages is very important in determining their phonological/phonetic behaviour. Interaction between the two languages should be interpreted in conjunction with whether it occurred when the bilinguals were communicating with monolinguals or other bilinguals. In the case of bilingual conversations, factors such as the base language, the degree of activation of each language, and the dominant language of the bilingual will all play a role in phonological patterns observed. When these factors are taken into consideration , it can be concluded that the bilinguals in this study did acquire separate production patterns for each of their languages in relation to the variables examined. In each language, the patterns were similar to those of the monolingual controls in the study when the productions occurred in the corresponding language sessions. During these sessions, signs of influence between the two languages were minimal and point to the bilinguals' overall ability to keep the phonologies of their languages separate. During the Arabic session with the mothers, the bilinguals used communicative strategies such as code-switching and borrowing in order to avoid dysfluency and to keep the communication going. Since these strategies were used only with interlocutors that also spoke and understood the two languages, the bilingual subjects can be said to have shown signs of sociolinguistic competence at a fine-grained phonetic level. This can also be seen in the way the bilinguals exhibited awareness of and adopted accent features that part of their community and that are undergoing change.

Language dominance in the bilinguals from this study certainly played a role in the type and amount of code-switching and borrowing that took place, especially that it was almost restricted to English code-switches during Arabic interactions. But as Grosjean

(1995: 259) points out, bilinguals have developed competencies in their languages to the extent required by their needs and those of the environment. Because the needs and uses of the languages are usually quite different, bilinguals are rarely equally or completely fluent in their languages. However, new situations, new environments, and new interlocutors will involve new linguistic needs and might induce a change in the competence of the bilingual, such as the acquisition of prevoicing by B7 following 18 months of attending a weekend Arabic school. These are, however, intricately related to the age and the complexity of the feature being acquired, as evidenced by the fact that B7's brother, B10, did not show similar signs of acquisition despite having experienced the same changes in input.

On the whole, the monolingual Arabic controls produced patterns that were similar to the ones normally described in the literature, but detailed auditory and instrumental analysis helped reveal important acoustic features in the sounds that were examined that have not been well-documented. These include the weak tap variant and the emphatic realization of glottal stops that were found for some of the Arabic speakers in the study. These and other phonetic features of Lebanese Arabic will be the subject of further investigation in the future. As for the monolingual English patterns, data from the subjects from this study and from Leeds speakers from the IViE corpus (Grabe & Nolan, 2001) point to the need for up-to-date phonological studies of the accents of Leeds and York and of the sociolinguistic patterning of variation in the respective communities. Chambers (2002b: 123) points out that, in order to obtain a representative sample of the populations for a survey of urban as well as rural areas, we need to include not only men and women of all classes and ages, but also residents of the survey area who are relative newcomers to it. This study suggests that surveys should also include both bilingual and monolingual speakers of the language, since both groups have been shown to be taking part in language variation and changes in their community.

PAGE
MISSING
IN
ORIGINAL

APPENDIX ONE

Questionnaires

Parents' Questionnaire

A. General information

1. How long have you lived in this country?
2. How old were you when you came over to the UK?
3. What were your reasons for coming to, and then living in the UK?
4. What languages do you speak?
5. What is your level of education?
6. What is your occupation?
7. Language use: what language(s) do you use with:
   - your spouse
   - your neighbours
   - your co-workers
   - your friends
   - your children

B. Social relationships

1. Do you have British and/or Arab friends in the UK? Who do you socialise with the most?
2. Are there other Lebanese families in Leeds/York that you are in contact with?
3. What nationality are your neighbours/ co-workers? How would you describe your relationships with them?
4. Are you satisfied with your life in the UK and do you feel that you and your family are well-integrated into society?
5. How often do you go back to the Lebanon?
6. Do you intend to reside in the UK permanently or do you have plans to go back to the Lebanon in the future?
7. How many children have you got? How old are they?

C. Information about each child

1. Where was your child born? Where has he/she been living since?
2. How often does your child go to the Lebanon and for how long?
3. What language(s) (in chronological order) was your child exposed to before nursery? What was the frequency of exposure to each language?
4. At what age was your child first heard speaking English/ Arabic?
5. What decisions did you make about which language(s) to use with your child? What were the reasons (e.g. cultural, religious...) behind your decisions?
6. At what age did your child start going to nursery/school?
7. Can you describe your child's linguistic development since he/she started attending nursery/school?
8. Did your child face any linguistic or social problems when he/she started attending nursery/school? How did the nursery/school cater for those problems? How did you cater for them?
9. Does your child attend any other school apart from the mainstream English-medium school? (There are certain part-time Arabic schools in Britain that are funded by Arab communities and that offer Arabic literacy to children at weekends). If yes, at what age did he/she start attending it and how do you feel this has affected their language development, choice, or dominance?

10. What language(s) does your child use with:
    - you
    - the neighbours
    - relatives
    - friends
    - brothers or sisters
11. What language(s) is your child exposed to when:
    - watching TV/videos
    - listening to music
    - reading (if applicable)
    - playing games
12. What is the literacy level of your child in each language?
13. Which language do you consider to be the dominant language of your child? Can you also rate him/her on each language according to the scale below?

1- Cannot speak the indicated language, has a few words or phrases, cannot produce sentences (Expressive Language), only understands a few words (Receptive Language).
2- Cannot speak the indicated language, has a few words or phrases (Expressive Language), understands the general idea of what is being said (Receptive Language).
3- Limited proficiency with grammatical errors, limited vocabulary (Expressive Language), understands the general idea of what is being said (Receptive Language).
4- Good proficiency with some grammatical errors, some social and academic vocabulary (Expressive Language), understands most of what is said (Receptive Language).
5- Native-like proficiency with few grammatical errors, good vocabulary (Expressive Language), understands most of what is said (Receptive Language).
DK- Don't Know.

14. Which language do you consider is your child's favourite?
15. Do you try to influence your child's language choice/use? How do you do that and why?
16. What is your opinion about your child growing up bilingual?
17. What is your opinion about your child growing up in the UK?

Teacher's questionnaire

1. How long have you known the subject?
2. Compared to his/her monolingual English peers, how would you rate the subject's overall ability in English, if possible in each of the skills, but most importantly with regard to his/her spoken ability and degree of fluency in the language? What, if anything, can you attribute to his/her bilingual background?

1- Cannot speak the indicated language, has a few words or phrases, cannot produce sentences (Expressive Language), only understands a few words (Receptive Language).
2- Cannot speak the indicated language, has a few words or phrases (Expressive Language), understands the general idea of what is being said (Receptive Language).
3- Limited proficiency with grammatical errors, limited vocabulary (Expressive Language), understands the general idea of what is being said (Receptive Language).
4- Good proficiency with some grammatical errors, some social and academic vocabulary (Expressive Language), understands most of what is said (Receptive Language).
5- Native-like proficiency with few grammatical errors, good vocabulary (Expressive Language), understands most of what is said (Receptive Language).
DK- Don't Know.

3. What type of test does the school use to assess the pupil's initial language ability/ proficiency for identification and placement? Is the test used later to monitor the pupil's progress?
4. Would you consider the subject's English accent as native-like? What (if any) are the features that you have noticed to be non-native?
5. Would you consider the subject's English accent as marked by any regional variation, for e.g. Leeds/York accent? If not, what comments do you have about his/her accent?
6. How would you describe the subject's interpersonal language skills?
7. How would you describe the subject's social interactions and relationships with his/her classmates? Do you have any comments about his/her degree of integration in the school society? What, if anything can you attribute to his/her ethnic minority background?
8. Are the subject's classmates mostly English monolinguals or are the students from a wide variety of language backgrounds (and thus possibly bilingual as well or speakers of English as a foreign language)?
9. From your intuition as a native speaker, can you tell whether the majority of monolingual English speakers in the subject's classroom are originally from Leeds/York (and thus have a Leeds/York accent) or from other cities in the UK (and thus have different accents)?
10. Is there any accent the school is likely to encourage in the students (through the general language use of the teachers/ administrators)?
11. Are there any comments you would like to add or other issues that you think can be further investigated?

Children's Questionnaire

1. How old are you?
2. How many languages do you speak?
3. Where did you learn each language?
4. Which language do you think you know better?
5. Which language do you prefer? Why?
6. What language(s) do you use with your:
   - parents
   - brother/sister (if applicable)
   - neighbours
   - friends
   - relatives
7. Which language(s) do you:
   - count in
   - think in
   - tell jokes in
   - swear in
   - dream in

   N.B. Younger subjects might find it hard to answer this question.
8. (For those who attend both English and Arabic schools) Which of the two schools do you like better? Why?
9. Do you like living in the UK?
10. Who are your friends in the UK and where do they come from? Who do you spend time with the most?
11. How often do you go back to the Lebanon?
12. What do you like and/or dislike about your life in Britain?
13. What do you like and/or dislike about the Lebanon?
14. Which country do you prefer and why? Where would you like to live?
15. (Where applicable) Which of the two schools do you prefer and why?

APPENDIX TWO

Sample data elicited from a picture-book

English	Arabic	English	Arabic
Animals	**ħajawe'ne:t**	Butter	'zɪbde
Hen	'dʒe:ʒe	Eggs	be:dˤ
Chickens	'sˤi:sa:n	Meat	'laħme
Peacock	'tˤa:wu:s	Tea	ʃa:j
Cockerel	di:k	Pot	tˤanʒara
Dog	'kaleb	Coffee	'ʔahwe
Butterfly	fa'ra:ʃe	Beans	ban
Worm	'du:de	Tin	'ʕɪlbe
Goat	'ʕanze	Water	maj
Lion	'ʔasad	Bottles	ʔa'ne:ne
Horse	ħsˤa:n	Beer	bi:ra
Bear	dɪb		
Cow	'baʔra	**Body parts**	**ʔaʕ'dˤa:ʔ 'ʒɪsɪm**
Grass	'ʕɪʃɪb	Stomach	batˤɪn
Giraffe	za'ra:fe	Nose	mɪn'xa:r
Camel	'ʒamal	Head	ra:s
Elephant	fi:l	Neck	'raʔbe
Whale	ħu:t	Mouth	tɪm
Shark	'ʔɪrɪʃ	Teeth	sne:n
		Muscles	ʕadˤa'le:t
Food	**'ʔakɪl**	Elbow	ku:ʕ
Cherries	'karaz	Fingers	ʔa'sˤa:bɪʕ
Bananas	mo:z	Nails	dˤa'fi:r
Peach	dɪ'r:a:ʔ	Foot	'ʔɪʒɪr
Raspberries	tu:t	Toes	'ʔɪsˤbaʕ 'ʔɪʒɪr
Figs	ti:n	Ankle	'ke:ħɪl
Grapes	'ʕɪnab	Blood	dam:
Pears	nʒa:sˤ	Hand	ʔi:d
Tomatoes	bana'dura		
Peppers	'flajfle	**Clothing/ accessories**	**'ʔalbɪse**
Cucumber	xja:r	Hat	bɪr'najtˤa
Onions	'basˤal	Glasses	ʕwaj'ne:t
Potatoes	ba'ta:ta	Bag	'ʃantˤa
Garlic	tu:m	Earrings	'ħalaʔ
Yoghurt	'laban	Ring	'xa:tɪm
Orange	laj'mu:n	Jumper	'kanze
Juice	ʕa'sˤi:r	Scarf	ʃe:l
Carrots	'ʒazar	Socks	kal'se:t
Nuts	bzu:'ra:t	Sandals	'sˤandˤal
Bread	'xɪbɪz	Boot	sˤʊ'b:a:t

APPENDIX THREE

Summary of the number and type of recordings that were conducted for the current study

GK = author

Subject	Interlocutor	Activity
B5	GK	Picture-naming + story-telling in English
	BF5	Picture-naming + story-telling in Arabic
	E5	Free-play
E5	GK	Picture-naming + story-telling in English
EF5	GK	Word list reading + story-telling + interview in English
EM5	GK	Word list reading + story-telling + interview in English
A5	GK	Picture-naming + story-telling in Arabic
AF5	GK	Word list reading + story-telling + interview in Arabic
AM5	GK	Word list reading + story-telling + interview in Arabic
BF5	GK	Word list reading + story-telling + interview in Arabic
	GK	Word list reading + story-telling in English
BM5	GK	Word list reading + story-telling + interview in Arabic
	GK	Word list reading + story-telling in English
B7	GK	Picture-naming + story-telling in English
	BF7	Picture-naming + story-telling in Arabic
	E7	Free-play
	B10	Free-play
E7	GK	Picture-naming + story-telling in English
EF7	GK	Word list reading + story-telling + interview in English
EM7	GK	Word list reading + story-telling + interview in English
A7	GK	Picture-naming + story-telling in Arabic
BF7	GK	Word list reading + story-telling + interview in Arabic
	GK	Word list reading + story-telling in English
BM7	GK	Word list reading + story-telling + interview inArabic
	GK	Word list reading + story-telling in English
B10	GK	Picture-naming + story-telling in English
	BF7	Picture-naming + story-telling in Arabic
	E10	Free-play
	B7	Free-play
E10	GK	Picture-naming + story-telling in English
EF10	GK	Word list reading + story-telling + interview in English
EM10	GK	Word list reading + story-telling + interview in English
A10	GK	Picture-naming + story-telling in Arabic
AF10	GK	Word list reading + story-telling + interview in Arabic
AM10	GK	Word list reading + story-telling + interview in Arabic

APPENDIX FOUR

Accent rating experiment

You will now listen to the recordings of 12 children and adults telling stories. Each recording will last for about 30s. As you listen to the stories:

a) Give each speaker a number on a scale from 1 to 4 ranging from:
1 = definitely Native **2 = probably Native**
3 = probably Non-Native **4 = definitely Non-Native**

b) If you can, explain your choice by referring to specific **Linguistic Features** (pronunciation, vocabulary use, etc.).

c) If you can, try to define each speaker's accent as narrowly as you can (e.g. northern/southern, Yorkshire/Lancashire, Manchester/Leeds)

		Definitely native	**Probably native**	***Probably non-*native**	***Definitely non-*native**
Speaker					
1.	Features? Accent?	**1**	**2**	**3**	**4**
2.	Features? Accent?	**1**	**2**	**3**	**4**
3.	Features? Accent?	**1**	**2**	**3**	**4**
4.	Features? Accent?	**1**	**2**	**3**	**4**
5.	Features? Accent?	**1**	**2**	**3**	**4**
6.	Features? Accent?	**1**	**2**	**3**	**4**

		Definitely native	Probably native	Probably non-native	Definitely non-native
7.	Features? Accent?	1	2	3	4
8.	Features? Accent?	1	2	3	4
9.	Features? Accent?	1	2	3	4
10.	Features? Accent?	1	2	3	4
11.	Features? Accent?	1	2	3	4
12.	Features? Accent?	1	2	3	4

APPENDIX FIVE

Results of the accent-rating experiment

	Native	
	Accent	**Comments**
EF5 (1)	- Southern (13)	- posh - middle class - correct pronunciation + grammar
	- Sussex (1)	- southern vowels e.g. [əʊ]
	- Southern RP (10)	- BBC Eng (2)
	- RP (1)	- very well pronounced - no final glottal stops - quite nasal
	- Midlands/Southern (1)	
	- Standard but not Southern (1)	
	Other comments	- no regional accent
		- precise pronunciation
		- no slang, full articulation of sounds

	Native	
	Accent	**Comments**
EF5 (2)	- Southern (5)	- no final glottal stops - very fluent- good use of vocabulary- excellent pronunciation- correct stress patterns.
	- South East (2)	- [ʋ] for /r/
	- London (1)	
	- London RP (1)	- very well pronounced
	- Southern RP (5)	- fluent + articulates all sounds (2) - high pitched
	- Northern RP (1)	
	- Northern (9)	- [a] vowels (4) - northern intonation - middle class - not very pronounced regional accent
	- Midlands (1)	
	Other comments	- clear, fluent, appealing intonation

	Native	
	Accent	**Comments**
EM5	- Northern (8)	- [ʌ] > [ʊ] (4)
	- Leeds (1)	- [ʌ] > [ʊ] - dark [ɫ]
	- Liverpool/ Birmingham (1)	
	- Midlands (3)	
	- North Midlands (1)	
	- Nottingham/Derby (1)	- rhythm
	- Leicester (1)	- [ʌ] > [ʊ]
	- Birmingham (2)	- diphthongs - long vowels
	- *Midlands/southern (1)*	- [ɔ] > [aʊ] - *fluent*
	- Norfolk (1)	
	- East Norfolk (1)	
	- Southern (3)	- difficult to tell - use of [ʌ] rather than [ʊ] - use of [aʊ]
	- South east (2)	
	- South west (1)	- use of [ɑ] vowel
	- Southern/West Country (1)	
	- South West (1)	
	- Bristol (1)	

	Native	
	Accent	**Comments**
EM7	- Northern (6)	- [a] vowels (2) - [ʌ] > [ʊ] - monophthongs - posh - HRT (high rising tone)
	- Leeds (1)	
	- Sheffield (2)	- northern vowels - [a] vowels
	- Yorkshire (5)	- [a] vowels (2) - [ʌ] > [ʊ]
	- Yorkshire or Lancashire (1)	
	- Lancashire (3)	- [ʌ] > [ʊ] - intonation - rhoticity
	- Manchester (5)	- [a] vowels (2) - vowels
	- Slightly Northern (1)	- exceptions: use of [əʊ]
	- Northern with some Southern aspects (2)	- definite article reduction (1) - posh or upper class (1)
	- Southern (1)	
	- London RP (1)	
	- RP (1)	- BBC Eng.
	Other comments	- fluent - good vocabulary - excellent pronunciation - good grammar

	Native	
	Accent	**Comments**
E5	- Southern (7)	
	- South west (1)	
	- South east (1)	[θ] > [f]
	- Northern (2)	
	- Yorkshire (1)	- pronunciation of 'doggie'
	- Geordie (2)	
	- Irish (1)	- long vowels
	- RP (1)	
	Other comments	- simple and monosyllabic
		- [θ] > [f]
		- [i] in 'doggy'
		- hesitation (3) + pauses
	Non-native	
	Accent	**Comments**
E5	- unspecified (1)	- different intonation - stressed /i/ of 'doggy'

	Native	
	Accent	**Comments**
E7	- Northern (8)	- /h/ dropping - [ə] in 'and' - certain vocabulary
	- Northern/ Leeds (2)	- /h/ dropping
	- Slightly Northern (1)	
	- Manchester (3)	- colloquial pronunciations - tense vowel in 'he'
	- Manchester/ Liverpool (1)	
	- Southern (4)	
	Other comments	- not as fluent as the others (4)
		- very hesitant, slow speech (3)
		- struggles with some words
E7	**Non-native**	
	Accent	**Comments**
	- Italy/Spain (1)	
	- unspecified (3)	

	Native	
	Accent	**Comments**
E10	- Northern (10)	- vowels + glottal stops - [a] vowels (5)
	- Slightly northern (1)	
	- Yorkshire (2)	- use of 'goes' and 'sees' - /h/ dropping
	- Lancashire (2)	- [a] vowels
	- Northern Manchester (1)	- Manchester vowels + rhythm
	- Manchester (2)	
	- Liverpool/Manchester (1)	
	- NW/ Liverpool (2)	- Liverpudlian vowels
	- Cheshire (1)	- [a] vowels
	- Midlands (4)	- /h/ dropping
	- London Black vernacular (1)	
	Other comments	- hesitant, but complex vocabulary + correct grammar
		- no hesitations or pauses, fluent (2).

	Native	
	Accent	**Comments**
B5	- Southern (13)	- use of [ʌ] rather than [ʊ] (2) - [ð] > [v] - diphthong in 'told'
	- Southern RP (1)	- use of [ʌ] rather than [ʊ] - RP vowels
	- London (1)	
	- South West (1)	
	- West Country (1)	- [aɪ] > [ɔɪ] e.g. 'side'
	- Northern (2)	- [ð] > [v]
	- Yorkshire (1)	
	- Manchester (1)	
	- Midlands (1)	- [ʌ] > [ʊ]
	Other comments	- [w] for /r/
		- child syntactic errors (4)
		- correct syntax
		- overgeneralisation e.g. 'bited'(4)

	Native	
	Accent	**Comments**
B7	- Northern (10)	- *glottal stops (5)* - [ʌ] > [ʊ] (2) - grammar + pronunciation - *difficult to tell*
	- North east (1)	
	- Lancashire (1)	
	- Liverpool (1)	- *glottal stops* - intonation
	- Midlands (2)	
	- Midlands but Standard (1)	
	- Southern (4)	- phonological features
	- London/Essex (1)	
	- RP (1)	- precise pronunciation
	Other comments	- not very fluent (4)
		- some pauses
		- good grammar + pronunciation
		- complex vocabulary
		- *syntactic* errors, but common ones (2)
	Non-native	
	Accent	**Comments**
B7	- Indian (2)	- slightly retroflex sounds
	- Asian (1)	- retroflex sounds
	- Unspecified (1)	- [ð] > [d]

	Native	
	Accent	**Comments**
B10	- Southern (10)	- Standard phonology + morphology
	- RP	
	- Northern (3)	
	- Northern/ Leeds	
	- Northern/ Lancashire	
	- Sheffield/ Manchester	- short vowels
	- Manchester (2)	- Monophthong [a:] in 'shouted'
	- South Yorkshire	
	- Midlands	
	Other comments	- good vocabulary
		- fluent
		- pronunciation of all sounds, e.g. [t] for [ʔ] in 'little' (2)
		- unusual syntactic errors
		- certain problems with pronunciation but generally okay

	Non-native	
	Accent	Comments
B10	- Indian (1)	- nasalised vowels - stuttered
	- Not British English (1)	- vocabulary
	- unspecified (3)	- certain expressions

	Non-native	
	Accent	Comments
BF5	- European (1)	
	- German (8)	- rhythm - pronunciation (3) - [v] > [f] - lack of native-like pauses
	- French (2)	- [ð] > [z]
	- Eastern European (1)	
	- Asian (1)	
	- Indian accent (2)	- intonation distinct (2) - unaspirated /t/s
	Other comments	- [ð] > [z] (4)
		- [v] > [f]
		- [θ] > [s]
		- trills
		- syllable-timed rhythm (3)
		- NN pronunciation (4)
		- pronunciation of all consonants
		- NN intonation (2)
		- vocab + grammar (3)
		- lots of hesitations and pauses (3)
		- slow speech
		- unclear

	Non-native	
	Accent	**Comments**
BM5	- European (1)	- aspirated final /t/s
	- French (1)	- pronunciation and rhythm
	- *Swedish (1)*	
	- Scandinavian (2)	
	- European/Middle Eastern (1)	
	- Arabic (1)	- vowel sounds
	- Asian (1)	- pronunciation
	- Indian (3)	- [ð] > [d] (2) - careful speed - retroflex vowels - retroflex sounds
	- West Indian (1)	- [o] vowels
	- Caribbean (1)	- [o] vowels - [ð] > [d]
	Other comments	- [ð] > [z]
		- [ð] > [d]
		- [o] vowels (2)
		- taps
		- final /t/ deletion
		- NN intonation (2) + stress (1)
		- NN vowels
		- NN accent
		- unusual rhythm
		- pronunciation of consonants with length
		- hesitant (2)
		- not fluent (2)
		- long pauses
		- slow + deliberate speech (2)
		- grammatical mistakes (4)
	Native	
	Accent	**Comments**
	- Northern (1)	

PAGE

NUMBERING

AS ORIGINAL

BIBLIOGRAPHY

Abu-Haidar, F. (1979). *A Study of the Spoken Arabic of Baskinta*. Leiden & London: The Royal Asiatic Society.

Adnan Zawaydeh, B. (1998). Gradient uvularisation spread in Ammani-Jordanian Arabic. In A. Benmamoun, M. Eid & N. Haeri (eds.). *Perspectives on Arabic Linguistics XI*. Amsterdam/Philadelphia: John Benjamins. pp. 117-141.

Agnihotri, R.K. (1979). *Processes of Assimilation: a Sociolinguistic Study of Sikh Children in Leeds*. Unpublished PhD dissertation, University of York.

Akin, J., Goldberg, A., Myers, G. and Stewart, J. (eds.) (1970). *Language Behaviour: a Book of Readings*. The Hague: Mouton.

Al Ani, S.H. (1970). *Arabic Phonology: An Acoustical and Physiological Investigation*. Paris: Mouton.

Albert, M.L. & Obler, L.K. (1978). *The Bilingual Brain: Neuropsychological and Neurolinguistic Aspects of Bilingualism*. New York: Academic Press.

Allen, G.D. (1985) How the young French child avoids the pre-voicing problem for word-initial voiced stops. *Journal of Child Language*, **12**: 37-46.

Allport, A. Mackay, D. Prinz, W. & Scheerer E. (eds.) (1989). *Language Perception and Production*. New York: Academic Press.

Al-Shareef, J. (2002). *Language Change and Variation in Palestine: a Case Study of Jabalia Refugee Camp*. Unpublished PhD dissertation, University of Leeds.

Amayreh, M.M. & Dyson, A.T. (1998). The acquisition of Arabic consonants. *Journal of Speech and Language Hearing Research*, **41**: 642-653.

_______ & _______ (2000). Phonetic inventories of young Arabic-speaking children. *Clinical Linguistics & Phonetics*, **14(3)**: 193-215.

Anani, M. (1985). Differences in the distribution between Arabic /l/, /r/, and English /l/, /r/. *Papers and Studies in Contrastive Linguistics*, **20**: 129-133.

Andersen, E.S. (1990). *Speaking with Style: the Sociolinguistic Skills of Children*. London & New York: Routledge.

Aslin, R.N. & Pisoni, D.B. (1980). Effects of early linguistic experience on speech discrimination by infants: a crtitique of Eilers, Gavin, & Wilson. *Child Development*, **51**: 107-112.

_______, _______, Hennessy, B.L. & Perey, A.J. (1981). Discrimination of voice onset time by human infants: new findings and implications for the effects of early experience. *Child Development*, **52**: 1135-1145.

Ball, M.J., Müller, N., & Munro, S. (2001a). The acquisition of rhotic consonants by Welsh-English bilingual children. *International Journal of Bilingualism*, **5(1)**: 71-86.

_______, _______, & _______ (2001b). Patterns in the acquisition of the Welsh lateral fricative. *Clinical Linguistics and Phonetics*, **15**, 3-7.

Bar-Adon, A. & Leopold, W. (eds.) (1971). *Child Language*. Englewood Cliffs, NJ: Prentice-Hall.

Bayley R. & Preston, D. (eds.) (1996). *Second Language Acquisition and Linguistic Variation*. Amsterdam: John Benjamins.

Bell, A. (1984). Language style as audience design. *Language in Society*, **13**: 145-204.

Benmamoun, A., Eid, M. & Haeri, N. (eds.) (1998) *Perspectives on Arabic Linguistics XI*. Amsterdam/Philadelphia: John Benjamins.

Berko Gleason, J. (ed.) (1994) *Language Development*. Columbus: Merrill.

Berman, R.A. & Slobin, D.I. (1994). *Relating Events in a Narrative: a Crosslinguistic Developmental Study*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bernhardt, B. & Stemberger, J. P. (1998). *Handbook of Phonological Development: from the Perspective of Constraint-based Non-linear Phonology*. San Diego, California: Academic Press.

Best, C.T. (1995). A direct realist view of cross-language speech perception. In W. Strange (ed.). *Speech Perception and Linguistic Experience: Issues in Cross-Linguistic Research.* Baltimore: York Press. pp. 171-206.

Bialystok, E. (1991) *Language Processing in Bilingual Children.* Cambridge: Cambridge University Press.

________ (ed.) (2001). *Bilingualism in Development: Language, Literacy, & Cognition.* Cambridge: Cambridge University Press.

Bladon, R.A. & Al-Bamerni, A. (1976). Coarticulation resistance in English /l/. *Journal of Phonetics,* 4: 137-150.

Bond, Z.S., Eddey, J.E, & Bermejo, J.J. (1980). VOT del espanol to English: comparison of a language-disordered and normal child. *Journal of Phonetics,* **8**. 287-290.

Borden, G.J. & Harris, K.S. (1984). *Speech Science Primer: Physiology, Acoustics, and Perception of Speech* (2nd ed.). Williams & Wilkins: London.

Bosch, L. & Sebastián-Gallés, N. (1997). Native language recognition abilities in 4-month-old infants from monolingual and bilingual environements. *Cognition,* **65**: 33-69.

_______ & ______ (2002). Early language differentiation in bilingual infants. In J. Cenoz & F. Genesee (eds.) *Trends in Bilingual Acquisition.* Amsterdam: John Benjamins.

Broe, M. and Pierrehumbert, J. (eds.) (2000). *Papers in Laboratory Phonology V: Language Acquisition and the Lexicon.* Cambridge: Cambridge University Press.

Browman, C.P. & Goldstein, L. (1986). Towards an articulatory phonology. *Phonology Yearbook, 3*: 219-252.

________ & Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica,* **49**: 155-180.

Brown, G. (1977). *Listening to Spoken English.* London: Longman.

Burling, R. (1971). Language development of a Garo and English speaking child. In A. Bar-Adon & W. Leopold (eds.). *Child Language.* Englewood Cliffs, NJ: Prentice-Hall. pp. 170-185.

Burton-Roberts, N., Carr, P. & Docherty, G. (eds.). (2000). *Phonological Knowledge: Conceptual and Empirical Issues.* Oxford: Oxford University Press.

Butterworth, B. Comrie, B. & Dahl, Ö. (eds.). *Explanations for Language Universals.* Berlin: Mouton.

Cantineau, J. (1960). *Etudes de Linguistique Arabe.* Paris: Librairie Clincksieck.

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 (2)**: 421-428.

Carter, P. (1999). Abstractness in phonology and extrinsic phonetic interpretation: the case of liquids in English. *Proceedings of the XIVth International Congress of Phonetic Sciences.* pp. 105-108.

Catford, J.C. (1988). *A Practical Introduction to Phonetics.* Oxford: Oxford University Press.

Cenoz J. & Genesee F. (eds.) (1998). *Beyond Bilingualism: Multilingualism and Multilingual Education.* Clevedon: Multilingual Matters.

________ & ________ (eds.) (2002) *Trends in Bilingual Acquisition.* Amsterdam: John Benjamins.

Chambers, J.K. (1973). Canadian raising. *Canadian Journal of Linguistics,* **18**: 113-135.

________ (1992). Dialect acquisition. *Language,* **68**: 673-705.

________ (2002a). *Sociolinguistic Theory* (2nd ed). Oxford: Blackwell.

________ (2002b). Dynamics of dialect convergence. *Journal of Sociolinguistics,* **6** (1): 117-130.

Champ, P. (1983). The evaporation of liquids in Cockney. *Nottingham Linguistic Circular,* **12(1)**: 1-20.

Chantraine, Y., Joanette, Y. & Ska, B. (1998) Conversational abilities in patients with right hemisphere damage. *Journal of Neurolinguistics,* **11**: 21-32.

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

Cho, T & Ladefoged, P. (1999). Variation and universal in VOT: evidence from 18 languages. *Journal of Phonetics*, **27**: 207-227.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

_______ & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper and Row.

Clyne, M. (1972). *Perspectives on Language Contact*. Melbourne: the Hawthorne Press.

Comas, M.T. (1986). *Caracteritzacio Acustica de Realisacions the [l] en Parlants Bilingue Catalan*. Unpublished dissertation, Departement de Filologie Catalana, Universite Autonoma de Barcelona.

Connell, B. (1995). Some articulatory characteristics of the tap. In J. Windsor Lewis (ed.), *Studies in General and English Phonetics in Honour of Professor J D O'Connor* London: Routledge Kegan Paul.

Contreras, H. & Saporta, S. (1970). Phonological development in the speech of a bilingual child. In J. Akin, A. Goldberg, G. Myers, and J. Stewart (eds.). *Language Behaviour: a book of readings*. The Hague: Mouton. pp. 280-294.

Cooper, W.E. & Aslin, R.N. (1990). Preference in infant-directed speech in the first month after birth. *Child Development*, **61**: 1584-95.

Cutler, A. Mehler, J., Norris, D. & Segui. J. (1989). Limits on bilingualism. *Nature* **340**: 229-230.

Cruttenden, A. (1978). Assimilation in child language and elsewhere. *Journal of Child Language*, **5**: 373-378.

_______ (1994/2001). *Gimson's Pronunciation of English* (5th & 6th eds.). London: Arnold.

Crystal, D. (1986). Prosodic development. In P. Fletcher & M. Garman (eds.). *Language Acquisition (2nd ed.)*. Cambridge: Cambridge University Press.

Davenport, M. & Hannahs, S.J. (1998). *Introducing Phonetics & Phonology*. London: Arnold.

Davis, S. (1995). Emphasis spread and Grounded Phonology. *Linguistic Inquiry*, **26**: 465-498.

De Bot, K. (2000). A bilingual production model: Levelt's 'speaking' model adapted. In L. Wei. (ed.). *The Bilingualism Reader*. London & New York: Routledge. pp. 420-442.

de Boysson-Bardies, B. & Vihman, M. (1991). Adaptation to language: evidence from babbling and first words in two languages. *Language*, **67**: 297-319.

________, Vihman M., Roug-Hellichius, L., Durand, C. Landberg, I. & Arao, F. (1992). Material evidence of infant selection from target language: a cross-linguistic phonetic study. In C.A. Ferguson, L. Menn & C. Stoel-Gammon (eds.). *Phonological Development: Models, Research, Implications*. Timonium, MD: York Press. pp. 369-392.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L. Paulesu, E. (1997). Anatomical variability in the cortical representation of first and second language, *Neuroreport*, **8**: 3809-3815.

De Houwer, A. (1990). *The Acquisition of Two Languages from Birth: a Case Study*. Cambridge: Cambridge University Press.

________ (1995). Bilingual language acquisition. In P. Fletcher & B. MacWhinney (Eds.), *Handbook of Child Language*. Oxford: Blackwell.

________ (1998). By way of introduction: methods in studies of bilingual first language acquisition. *The International Journal of Bilingualism*, 2(3): 249-263.

De Groot, A.M.B. (1993). Word-type effects in bilingual processing tasks: support for a mixed representational system. In R. Schreuder & B. Weltens (eds.). *The Bilingual Lexicon*. Amsterdam: John Benjamins. pp. 27-51.

Delattre, P. (1971). Pharyngeal features in the consonants of Arabic, German, Spanish, French, & American English. *Phonetica*, **23**: 129-155.

Deuchar, M. & Clark, A. (1995). Early bilingual acquisition of the voicing contrast in English and Spanish. *Bangor Research Papers in Linguistics,* **6**: 24-37.

________ & Quay, S. (1999). Language choice in the earliest utterances: a case study with methodological implications. *Journal of Child Language,* **26** (2): 461-475.

________ & ______ (2000). *Bilingual Acquisition: Theoretical Implications of a Ccase Study.* Oxford: Oxford University Press.

Dickins, J. (ed.) (1996). *Three Topics in Arabic Phonology.* Centre of Middle Eastern and Islamic Studies, Occasional Papers **53**, University of Durham.

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

________ & Foulkes, P. (1999). Derby and Newcastle: instrumental phonetics and variationist studies. In P. Foulkes & G. Docherty (eds.) *Urban Voices.* London: Arnold. pp.47-71.

________ & _______ (2000). Speaker, speech and knowledge of sounds. In N. Burton-Roberts, P. Carr & G.J. Docherty (eds.) *Phonological Knowledge. Conceptual and Empirical Issues.* Oxford: Oxford University Press. pp. 105-129.

________, _______, Milroy J., Milroy L. & Walshaw, D. (1997). Descriptive adequacy in phonology: a variationist perspective. *Journal of Linguistics,* **33**: 275-310.

________, _______, Tillotson J. & Watt D. (2002). On the scope of phonological learning: issues arising from socially structured variation. Paper presented at the Labphon 8 Conference, Yale University, June 2002.

Dopke, S. (1992). *One Parent One Language: an Interactional Approach.* Amsterdam: John Benjamins.

Durkin, K. (ed.) (1986). *Language Development in the School Years.* Cambridge MA: Brookline Books.

Dyson, A.T. & Amayreh, M.M. (2000). Phonological errors and sound changes in Arabic-speaking children. *Clinical Linguistics & Phonetics,* **14(2)**: 79-109.

Eckert, P. (2000). *Linguistic Variation as Social Practice.* Oxford: Blackwell.

Edwards, M. L. (1973). The acquisition of liquids. *Ohio State University Working Papers in Linguistics,* **15**: 1-54.

_____ (1974). Perception and production in child phonology: the testing of four hypotheses. *Journal of Child language,* **1**: 205-19.

_____ (1978). *Patterns and Processes in Fricative Acquisition: Longitudinal Evidence from Six English Learning Children.* Doctoral dissertation, Stanford University.

Eilers, R.E., Gavin, W.J. & Oller, D.K. (1981). Cross-linguistic perception in infancy: early effects of linguistic experience. *Journal of Child Language,* **9**: 289-302.

_____, Oller, D.K & Benito-Garcia, C.R. (1984). The acquisition of voicing contrast in Spanish and English learning infants and children: a longitudinal study. *Journal of Child Language,* 11: 313-336.

_____, _____, & Lavoie, L. (1985). Where's the drift in babbling drift? A cross-linguistic study. *Applied Psycholinguistics,* **6**: 3-15.

Eimas, P.D., Miller, J.L. & Jusczyk, P.W. (1987). On infant perception and the acquisition of language. In S. Harnad (ed.). *Categorical Perception.* New York: Cambridge University Press.

Ellis A.W. (ed.) (1982). *Normality and Pathology in Cognitive Functions.* New York & London: Academic Press.

Elman, J. Diehl, R. & Buchwald, S. (1977). Perceptual switching in bilinguals. *Journal of the Acoustical Society of America,* **62**: 971-974.

Enstrom, D.H. (1982). Infant labial, apical, and velar stop productions: a voice onset time analysis. *Phonetica,* **39**: 47-60.

Fabbro, F. (1999). *The Neurolinguistics of Bilingualism.* Hove: Psychology Press.

Fantini, A.F. (1985). *The Language Acquisition of a Bilingual Child.* Clevedon: Multilingual Matters.

Ferguson, C.A. (1956). The emphatic *l* in Arabic. *Language,* **23**(3): 446-552.

_______ (1976). Remarks on theories of phonological development. In W. von Raffler-Engel & Y. Leburn (eds.) *Baby Talk and Infant Speech.* Amsterdam: Swets & Zeitlinger. pp. 84-98.

_______ & Farwell, C.B. (1975). Words and sounds in early language acquisition. *Language*, **51**: 419-439.

Ferguson, C.A. Menn L. & Stoel-Gammon C. (eds.) (1992). *Phonological Development: models, research, implications.* Timonium, MD: York Press.

Flege, J.E. (1980). Phonetic approximation in second language acquisition. *Language Learning*, **30**(1): 117-134.

_______ (1987). The production of 'new' and 'similar' phones in a foreign language: evidence for the effect of equivalence classification. *Journal of Phonetics*, **15**: 47-65.

_______ (1995). Second language speech learning: theory, findings, and problems. In W. Strange (ed.). *Speech Perception and Linguistic Experience: theoretical and methodological issues.* Baltimore: York Press. pp. 233-272.

_______ & Eefting, W. (1987). Production and perception of English stops by native Spanish speakers. *Journal of Phonetics*, **15**: 67-83.

_______ & Hammond, R.M. (1982). Mimicry of Non-distinctive phonetic differences between language varieties. *Studies in Second Language Acquisition*, **5(1)**: 1-17.

_______ Munro, M.J. & Mackay, I.R.A. (1996). Factors affecting the production of word-initial consonants in a second language. In R. Bayley & D. Preston (Eds.) *Second Language Acquisition and Linguistic Variation.* Amsterdam: John Benjamins. pp. 47-73.

_______ & Port, R. (1981). Cross-language phonetic interference: Arabic to English. *Language & Speech*, **24** (2): 125-145.

_______ & Slobin, D.A. (eds.) (1973). *Studies of Child Language Development.* New York: Holt, Rinehart and Winston.

Fletcher P. & Garman M. (eds.) (1979). *Language Acquisition.* Cambridge: Cambridge University Press.

_______ & _______ (eds.) (1986). *Language Acquisition* (2nd ed.) Cambridge: Cambridge University Press.

Fokes, J., Bond, Z.S. & Steinberg, M. (1985). Acquisition of the English voicing contrast by Arab children. *Language and Speech*, **28(1)**: 81-92.

Foulkes, P. (1997a). English [r]-sandhi - a sociolinguistic perspective. *Histoire, Epistémologie, Language*, **19**: 73-96.

_______ (1997b). Rule inversion in a British English dialect - a sociolinguistic investigation of [r]-sandhi in Newcastle upon Tyne. *University of Pennsylvania Working Papers in Linguistics*, **4(1)**: 259-270.

_______ & Docherty, G.J. (eds.) (1999) *Urban Voices.* London: Arnold.

_______ & _______ (2000). Another chapter in the story of /r/: 'labiodental' variants in British English. *Journal of Sociolinguistics*, **4**: 30-59.

_______, _______, & Watt. D. (1999). Tracking the emergence of sociophonetic variation: realisations of (t) by Newcastle children. *Leeds Working Papers in Linguistics and Phonetics*, **7**: 1-25. University of Leeds.

Fourcin, A. (1978). Acoustic patterns and speech acquisition. In N. Waterson & C. Snow (eds.). *The Development of Communication.* London: Wiley. pp. 144-171.

Fromkin, V. (ed.) (1985). *Phonetic Linguistics.* Orlando: Academic Press.

Gairdner, W.H.T. (1925). *The Phonetics of Arabic.* Oxford: Oxford University Press.

Gardner-Chloros, P. (1995). Code-switching in community, regional, and national repertoires: the myth of the discreteness of the systems. In L. Milroy & P. Muysken (Eds.) *One Speaker, Two Languages: Cross-disciplinary Perspectives on Codeswitching.* Cambridge: Cambridge University Press. pp. 68-89.

Garrett, M. (1982). Production of speech: observations from normal and pathological language use. In A.W. Ellis (ed.). *Normality and Pathology in Cognitive Functions.* New York & London: Academic Press. pp. 19-76.

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

________ (2001). Bilingual first language acquisition: exploring the limits of the language faculty. *Annual Review of Applied Linguistics*, **21**: 153-168.

________, Nicoladis, E. & Paradis, J. (1995). Language differentiation in early bilingual development. *Journal of Child Language*, **22**: 611-31.

Gibbon, F. (1999). Undifferentiated lingual gestures in children with articulation/ phonological disorders. *Journal of Speech, Language & Hearing Research*, **42**: 382-397.

Giddens, A (1989). *Sociology* (2nd ed.). Cambridge: Polity Press.

Gilbert, J.H.V. (1977). A voice onset time analysis of apical stop production in three-year-olds. *Journal of Child Language*, **4**: 103-110.

Goldinger, S. (1997). Speech perception and production in an episodic lexicon. In K. Johnson & J.W. Mullennix, (eds.). *Talker Variability in Speech Processing*. San Diego: Academic Press. pp. 33-66.

_______ (1998) Echoes of echoes? An episodic theory of lexical access. *PsychologicalReview*, 105: 251-279.

_______, Pisoni, D.B. & Logan, J.S. (1991). On the locus of talker variability effects in recall of spoken word lists. *Journal of Experimental Psychology: learning, memory, and cognition*, **17**: 152-162.

Goodz, N.S. (1989). Parental language mixing in bilingual families. *Journal of Infant Mental Health*, **10**: 25-44.

Gorlitzer von Mundy, V. (1983 [1959]). A 94-year-old with one German language center and probably two Slovenian centers. In M. Paradis (ed.), *Readings on Aphasia in Bilinguals and Polyglots*. Montreal: Didier. pp. 624-625.

Grabe, E. & Nolan, F. (2001). *English Intonation in the British Isles. The IViE Corpus.* CD ROMs produced as part of ESRC grant R000237145.

________, Post, B. & Watson, I. (1999). The acquisition of rhythmic patterns in English and French. *Proceedings of the XIVth International Congress of Phonetic Sciences*.

Green, D. (2000). Control, activation, and resource: a framework and a model for the control of speech in bilinguals. In W. Li (ed.). *The Bilingualism Reader*. London & New York: Routledge. pp. 407-419.

Greenberg J.H. (ed.) (1978). *Universals of Human Language*. Stanford, CA: Stanford University Press.

Grosjean, F. (1982). *Life with Two Languages: An Introduction to Bilingualism.* Cambridge, MA: Harvard University Press.

________ (1995). A psycholinguistic approach to codeswitching: The recognition of guest words by bilinguals. In L. Milroy & P. Muysken (eds.), *One Speaker, Two Languages: Cross-disciplinary Perspectives on Codeswitching*. Cambridge: Cambridge University Press. pp. 259-275.

________ (1998). Studying bilinguals: methodological and conceptual issues. *Bilingualism: Language and Cognition*, **2**: 131-149.

________ (2000). Processing mixed language: issues, findings, and models.. In W. Li. (ed.). *The Bilingualism Reader*. London & New York: Routledge. pp. 443-469.

________ (2001). The bilingual's language modes. In J.L. Nicol (ed.) *One Mind, Two Languages*. Oxford: Blackwell. pp. 1-22.

________ & Miller, J. (1994). Going in and out of languages: an example of bilingual flexibility. *Psychological Science*, **5**: 201-6.

Grunwell, P. (1986). Aspects of phonological development in later childhood. In K. Durkin, (ed.) *Language Development in the School Years*. Cambridge MA: Brookline Books.

Guasp, F. P. (2001). Catalan dark /l/ in Palma: the social interpretation of a change in apparent time. *Proceedings of the First International Conference on Language Variation in Europe*. Barcelona: Universitat Pompeu Fabra. pp. 161-167.

Haddad, G, (1984). *Problems and Issues in the Phonology of Lebanese Arabic*. Doctoral Dissertation, University of Illinois, Urbana-Champaign.

Hamers, J.F. & Blanc, M.H.A. (2000). *Bilinguality and Bilingualism* (2nd ed.). Cambridge: Cambridge University Press.

Hardcastle, W.J. & Barry W.J (1985). Articulatory and perceptual factors in /l/ vocalization in English. *Work in Progress*, **5**: 31-44. University of Reading.

________ & Laver, J. (eds.) (1997). *The Handbook of Phonetic Sciences*. Oxford: Blackwell.

________ & Marchal, A. (eds.) (1990). *Speech Production and Speech Modelling*. Dordrecht: Kluwer.

Harrel, R. (1957). *The Phonology of Colloquial Egyptian Arabic*. New York: American Council of Learned Societies.

Harnad S. (ed.) (1987). *Categorical Perception*. New York: Cambridge University Press

Haugen, E. (1956). *Bilingualism in the Americas: a Bibliography and Research Guide*. Montgomery, AL: University of Alabama Press.

______ (1977). Norm and deviation in bilingual communities. In P. Hornby (ed.). *Bilingualism: Psychological, Social, and Educational Implications*. New York: Academic Press.

Hazan, V. & Barrett, S. (2000). The development of phonemic awareness in children aged 6-12. *Journal of Phonetics*, **28**: 377-396.

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

Hecht, B. and Mulford, R. (1982). The acquisition of a second language phonology. Interaction of transfer and developmental factors. *Applied Psaycholinguistics*, **3**: 313-328.

Heselwood, B. (1996). Glottal states and emphasis in Baghdadi and Cairene Arabic: synchronic and diachronic aspects. In Dickins, J. (ed.) *Three Topics in Arabic Phonology*. Centre of Middle Eastern and Islamic Studies, Occasional Papers, **53**, University of Durham. pp. 20-44.

________ (1998). An unusual kind of sonority and its implications for phonetic theory. *Leeds Working Papers in Linguistics & Phonetics*, **6**: 66-80.

________ & McChrystal, L. (1999). The effect of age-group and place of L1 acquisition on the realisation of Panjabi stop consonants in Bradford: an acoustic sociophonetic study. *Leeds Working Papers in Linguistics & Phonetics* **7**: 49-68.

________ & ________ (2000). Gender, accent features and voicing in Punjabi-English bilingual children. *Leeds Working Papers in Linguistics & Phonetics*, **8**: 45-70.

Hewlett, N., Scobbie, J.M. & Matthews, B. (1999). Vowel duration in Scottish English speaking children. *Proceedings of the XIVth International Congress of Phonetic Sciences*. University of California, Berkeley. *pp. 2157-2160*.

Hintzman, D. L. (1986). '"Schema abstraction" in a multiple-trace memory model'. *Psychological Review*, **93**: 411-428.

Hiroko, K. (1998). Determinants of language choice in bilingual children: the role of input. *The International Journal of Bilingualism*, 2 (3): 327-346.

Hirschfeld, L.A. & Gelman, S.A (1997). What children think about language variation and social development. *Cognitive Development*, **12**: 213-238.

Hoffmann, C. (1991). *An Introduction to Bilingualism*. New York: Longman.

Holm, A. & Dodd, B. (1999). A longitudinal study of the phonological development of two Cantonese-English bilingual children. *Applied Psycholinguistics*, **20**: 349-376.

Holmes, J. (1992). *An Introduction to Sociolinguistics*. London: Longman.

Hornby, P. (ed.) (1977). *Bilingualism: Psychological, Social, and Educational Implications*. New York: Academic Press.

Huertos, F., Sabio, E., Silvestre, B. & Sonia, J. (1988). *Caracterisacion Acustica de /l/ Castellana en Catalano-Hablantes de Gerona en Habla Continua.* Unpublished paper, Laboratori de Fonetica, UAB.

Hughes, A. & Trudgill, P.J. (1996). *English Accents and Dialects. An Introduction to Social and Regional Varieties of British English* (3rd ed.). London: Arnold.

Hussein, R. (1980). *The Case of Triglossia in Arabic: with Special Emphasis in Jordan.* Unpublished PhD Dissertation. University of New York at Buffalo.

Hyltenstam, K. & Obler L. (eds) (1989). *Bilingualism Across the Lifespan: Aspects of Acquisition, Maturity and Loss.* Cambridge: Cambridge University Press.

Hymes, D. (1974). *Foundations in Sociolinguistics: an Eethnographic Approach.* Philadelphia: University of Pennsylvania Press.

Ibn Khaldun. (1958). *Muqaddimma* (tr. Rosenthal, F.). London.

Ingram, D. (1976). *Phonological Disability in Children.* New York: Elsevier.

________ (1979). Phonological patterns in the speech of young children. In P. Fletcher & M. Garman (eds.) *Language Acquisition.* Cambridge: Cambridge University Press.

________ (1981/2). The emerging phonological system of an Italian-English bilingual child. *Journal of Italian Linguistics,* 1981/2: 95–113.

________ (1986). Phonological development: production. In P. Fletcher & M. Garman (eds.) *Language Acquisition* (2nd ed.) Cambridge: Cambridge University Press. pp. 223–239.

________ (1989). *First Language Acquisition: Method, Description and Explanation.* Cambridge: Cambridge University Press.

Jacobson, R. (ed.) (1997). *Code-switching Worldwide.* Berlin: Mouton de Gruyter.

Jakobson, R. (1941/68). *Child Language Aphasia and Phonological Universals.* The Hague: Mouton.

________ (1957). Muffaxxama, the 'emphatic' phonemes in Arabic. In E. Pulgram (ed.). *Studies Presented to Joshua Whatmough.* The Hague: Mouton. pp. 105-115.

_______ & Halle (1956). *Fundamentals of Language.* The Hague: Mouton.

Jesry, M. *(1996). Some Cognitively Controlled Coarticulatory Effects in Arabic and English, with Particular Reference to VOT.* Unpublished PhD dissertation. University of Essex.

Johnson, C.E. & Lancaster, P. (1998). The development of more than one phonology: a case study of a Norwegian-English bilingual child. *The International Journal of Bilingualism,* **2**(3): 265-300.

Johnson, K. (1997). Speech perception without speaker normalization. In Johnson, K. & Mullennix, J.W. (eds.) *Talker Variability in Speech Processing.* San Diego: Academic Press. pp. 145-165.

Johnson, K. & Mullenix, J.W. (eds.) (1997). *Talker Variability in Speech Processing.* San Diego: Academic Press.

Jones, D. (1972). *An Outline of English Phonetics (9th ed.).* Cambridge: Cambridge University Press.

Jusczyk, P.W. (1993). From general to language-specific capacities: the WRAP-SA model of how speech perception develops. *Journal of Phonetics,* **21**: 3-28.

________, Hohne, E.A., & Mandel, D.R. (1995). Picking up regularities in the structure of the native language. In W. Strange (ed.). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues.* Baltimore: York Press. pp. 91-120.

Kahn, M. (1975). Arabic emphatics: The evidence for cultural determinants of phonetic sex-typing. *Phonetica,* **31(1)**: 38-50.

Keating, P.A. (1984). Phonetic and phonological representations of stop consonant voicing. *Language,* **60**: 286-319.

________ (1990). The window model of coarticulation: articulatory evidence. In J. Kingston & M. Beckman (eds.). *Papers in Laboratory Phonology I: Between the*

Grammar and the Physics of Speech. Cambridge: Cambridge University Press. pp. 451-470.

_______ (ed.) (1994) *Phonological Structure and Phonetic Form. Papers in Laboratory Phonology 3*. Cambridge: Cambridge University Press.

_______, Linker, W., & Huffman, M. (1983). Closure duration of stop consonants. *Journal of Phonetics*, 11: 277-290.

_______, Mikoś, M.J. & Ganong, W.F. (1981). A cross-language study of range of voice onset time in the perception of initial stop voicing. *Journal of the Acoustical Society of America*, **70(5)**: 1262-1271.

Kenstowicz, M. (1994). *Phonology in Generative Grammar*. Cambridge, MA: MIT Press.

Kent, R., & Read, C. (1992). *The Acoustic Analysis of Speech*. San Diego: Singular.

Kerswill, P.E. (1994). *Dialects Converging: Rural Speech in Urban Norway*. Oxford: Clarendon Press.

_______ (1996). Children, adolescents and language change. *Language Variation and Change*, **8(2)**: 177-202.

_______ & Williams (2000). Creating a new town koine: children and language change in Milton Keynes. *Language and Society*, **29**: 65-115.

Kewley-Port, D. & Preston, M.S. (1974). Early apical stop production: a voice onset time analysis. *Journal of Phonetics*, 2: 195-210.

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

Kingston, J. & Beckman, M. (eds.) (1990). *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*. Cambridge: Cambridge University Press.

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

Kohler, K. (1984). Phonetic explanation in phonology: the feature fortis/lenis. *Phonetica*, **41**: 150-174.

Konefal, J.A. & Fokes, J. (1981). Voice onset time: the development of Spanish-English distinction in normal and language disordered children. *Journal of Phonetics*, **9**: 437-444.

Kravin, H. (1992). Erosion of a language in bilingual development. *Journal of Multilingual and Multicultural Development*, **13(4)**: 307-325.

Kuhl, P. (1993). Infant speech perception: a window on psycholinguistic development. *International Journal of Psycholinguistics*, **9**: 33-56.

______ (1994). Learning and representation in speech and language. *Current Opinion in Neurobiology*, 4: 812-822.

______ & Iverson, P. (1995). Linguistic experience and the 'Perceptual Magnet Effect'. In W. Strange (ed.). *Speech Perception and Linguistic Experience: theoretical and methodological issues*. Baltimore: York Press. pp. 121-154.

Labov, W. (1963). The social motivation of a sound change. *Word*, **19**: 273-309.

________ (1966). *The Social Stratification of English in New York City*. Washington DC: Center for Applied Linguistics.

________ (1972). *Sociolinguistic Patterns*. Oxford: Blackwell.

________ (ed.) (1980). *Locating Language in Time and Space*. New York: Academic Press.

________ (1994). *Principles of Linguistic Change. Vol. I: Internal Factors*. Oxford: Blackwell.

________ (2001). *Principles of Linguistic Change. Vol. II: Social Factors*. Oxford: Blackwell.

Ladefoged, P. (1982/1993). *A Course in Phonetics* (2nd & 3rd eds.). New York: Harcourt Brace Jovanovich.

________ & Maddieson, I. (1996). *The Sounds of the World's Languages*. Oxford: Blackwell.

Lanza, E. (1992). Can bilingual two-year-olds code-switch? *Journal of Child Language*, **19**: 633-58.

______ (1997). *Language Mixing in Infant Bilingualism: a Sociolinguistic Perspective.* Oxford: Oxford University Press.

Laufer, A. & Baer, T. (1988). The emphatic and pharyngeal sounds in Hebrew and Arabic. *Language and Speech*, **31 (2)**, 181-205.

Lavoie, L.M. (2001). *Consonant Strength: Phonological Patterns and Phonetic Manifestations.* New York and London: Garland Publishing.

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

Leather, J. & Allan, J. (eds.) (1992). *New Sounds. Proceedings of the 1992 Symposium on the Acquisition of Second-Language Speech.* Amsterdam: University of Amsterdam.

Lehiste, I. (1970). *Suprasegmentals.* Cambridge, MA: MIT Press.

Lehmann W.P. & Malkiel Y. (eds.) (1968). *Directions for Historical Linguistics.* Austin, Texas: University of Texas Press.

Lehn, W. (1963). Emphasis in Cairo Arabic. *Language*, **39**: 29-39.

Leopold, W.F. (1941/1970). *Speech Development of a Bilingual Child: a Linguist's Record* (Vol. 2). New York: AMS Press.

Levelt, W. (1989). *Speaking: from Intention to Articulation.* Cambridge, MA: MIT Press.

Li, W. (ed.) (2000). *The Bilingualism Reader.* London & New York: Routledge.

Lindau, M. (1985). The story of /r/. In V. Fromkin (ed.). *Phonetic Linguistics.* Orlando: Academic Press. pp. 157-168.

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W.J. Hardcastle & A. Marchal (eds.) *Speech Production and Speech Modelling.* Dordrecht: Kluwer. pp. 403-439.

_______ (1992). Phonological units as adaptive emergents of lexical development. In C.A. Ferguson, L. Menn & C. Stoel-Gammon (eds.). *Phonological Development: Models, Research, Implications.* Timonium, MD: York Press.

_______, MacNeilage, P.F. & Studdert-Kennedy, M. (1984). Self-organising processes and the explanation of phonological universals. In B. Butterworth, B. Comrie, & Ö Dahl (eds.). *Explanations for Language Universals.* Berlin: Mouton. PP. 181-203.

Lindholm, K. & Padilla, A. (1977). Language mixing in bilingual children. *Child Language*, **5**: 327-335.

Lisker, L. (1975). Is it VOT or a first-formant transition detector? *The Journal of the Acoustical Society of America*, **57**(6): 1547-1551.

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

______ & ______ (1967). Some effects of context on voice onset time in English stops. *Language & Speech* **10**. 1-27

_______ & _______ (1971). Distinctive features and laryngeal control. *Language*, **47** (4): 767-783.

Llamas, M-C. (1998). Language variation and innovation in Middlesbrough: a pilot study. *Leeds Working Papers in Linguistics and Phonetics*, **6**: 98-115.

Local, J. (1983). How many vowels in a vowel? *Journal of Child Language*, **10**: 449-53.

Locke, J.L. (1983). *Phonological Acquisition and Change.* New York: Academic Press.

Lyon, J. (1996). *Becoming Bilingual.* Clevedon: Multilingual Matters.

MacKain, K.S. & Stern, D.N. (1985). The concept of experience in speech development. In K.E. Nelson (ed.). *Children's Language, 5.* Hillsdale, NJ: Laurence Erlbaum. pp. 1-33.

Macken, M. (1980). Aspects of the acquisition of stop systems: a cross-linguistic perspective. In G. Yeni-Komshian, J.F. Kavanagh & C.A. Ferguson (eds.). *Child Phonology. Volume 1: Production.* New York: Academic Press.

_______ & Barton, D. (1979). The acquisition of the voicing contrast in English: a study of voice onset time in word-initial stop consonants. *Journal of Child Language*, **7**: 41-74.

_______ & _______ (1980) The acquisition of the voicing contrast in Spanish: a phonetic and phonological study of word-initial stop consonants. *Journal of Child Language,* **7**: 433-458.

Mackey, W.F. (1965). Bilingual interference: its analysis and measurement. *Journal of Communication*, **15**: 239-49.

Mackey W.F. & Andersson T. (eds.) (1977). *Bilingualism in Early Childhood.* Rowley, MA: Newbury House.

MacLaughlin, B. (1984). *Second Language Acquisition in Childhood (Vol. 1): Preschool Children.* Hillsdale, NJ: Lawrence Erlbaum.

MacNeilage, P. F. (1980). The control of speech production. In G.H. Yeni-Komshian, J.F. Kavanagh and C.A. Ferguson (eds.). *Child Phonology. Volume 1: Production,* New York: Academic Press. pp. 9-21.

_______ (ed.) (1983). *The Production of Speech.* New York: Springer-Verlag.

Major, R. (1977). Phonological differentiation in a bilingual child. *Ohio State Working Papers in Linguistics*, **22**: 88-122.

_______ (1992). Losing English as a first language. *Modern Language Journal*, **76**: 190-208.

Martin, C.S., Mullennix, J.W., Pisoni, D.B. & Summers, W.V. (1989). Effects of talker variability on recall of spoken word lists. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**: 676-684.

Martinez, G. (1989). Caracterizacion acustica de la consonante lateral alveolar en un corpus de habla continua realizacion por hablante bilingue. Paper submitted to the XVIII Simposio de la Sociedad Española de Linguistica, Madrid.

Martinez-Dauden, G. & Llisteri, J. (1990). Phonetic interference in bilingual speakers learning a third language: the production of lateral consonants. Paper presented at the Meeting of the 9th World Congress of Applied Linguistics, Thessaloniki, Greece. April, 15-21.

Matthews, B. (2001). *On Variability and the Acquisition of Vowels in Normally Developing Scottish Children (18-36 Months).* Unpublished PhD Thesis, Queen Margaret University College, Edinburgh.

Mayer, M. (1969). *Frog, where are you?* New York: A Puffin Pied Piper.

McCarthy, J (1994). The phonetics and phonology of Semitic pharyngeals. In P. Keating (ed.) *Phonological Structure and Phonetic Form. Papers in Laboratory Phonology 3.* Cambridge: Cambridge University Press. pp. 191-233.

McClelland, J. & Elman, J. (1986). The TRACE model of speech perception. *Cognitive Psychology*, **18**: 1-86.

Mehler, J., Jusczyk, P.W., Lambertz, G., Halsted, N., Betoncini, J. (1988). A precursor of language acquisition in young infants. *Cognition* 29, 144-178.

Meisel, J.M. (1989). Early differentiation of languages in bilingual children. In K. Hyltenstam, & L. Obler (eds), *Bilingualism Across the Lifespan: Aspects of Acquisition, Maturity and Loss.* Cambridge: Cambridge University Press. pp. 13-40.

Meltzoff, A.N. & Moore, M.K. (1983). Newborn infants imitate adult facial gestures. *Child Development*, **54**: 702-709.

Menn, L. (1971). Phonotactic rules in beginning speech. *Lingua*, **26**: 225-251.

_______ & Stoel-Gammon, C. (1994). Phonological Development. In J. Berko Gleason (ed.), *Language Development.* Columbus: Merrill. pp. 65-113.

Menyuk, P. (1971). *The Acquisition and Development of Language.* London: Prentice Hall.

Milroy L. & Muysken P. (eds.) (1995). *One Speaker, Two Languages: Cross-disciplinary Perspectives on Codeswitching.* Cambridge: Cambridge University Press.

Mitchell, T.F. (1993). *Pronouncing Arabic 2.* Oxford: Clarendon Press.

Mitleb, F. (1984a). Voicing effect on vowel duration is not an absolute universal. *Journal of Phonetics*, **12(1)**: 23-27.

_______ (1984b). Vowel length contrast in Arabic and English: a spectrographic test. *Journal of Phonetics*, **12**(3): 229-235.

Moon, C., Cooper, R.P., Fifeer, W.P. (1993). Two-day old infants prefer native language. *Infant Behaviour and Development,* 16: 495-500.

Moskowitz, A.I. (1970). The two-year-old stage in the acquisition of English phonology. *Language*, 46: 426-441.

Mullenix, J.W., Pisoni, D.B. & Martin, C.S. (1989). Some effects of talker variability on spoken word recognition. *Journal of the Acoustical Society of America*, **85**: 365-378.

Muysken, P. (1995). Code-switching and Grammatical Theory. In L. Milroy & P. Muysken (Eds.), *One Speaker, Two Languages: Cross-disciplinary Perspectives on Codeswitching*. Cambridge: Cambridge University Press. pp. 177-198.

Myers-Scotton, C. (1997). *Duelling Languages: Grammatical Structure in Code-Switching*. Oxford: Oxford University Press.

_______ & Jake (2001). *Bilingualism: Multiple Voices*. Oxford: Blackwell.

Nasr, R.T. (1966). *Colloquial Arabic: an Oral Approach*. Beirut: Librarie du Liban.

Nathan, L., Wells, B. & Donlan, C. (1998). Children's comprehension of unfamiliar regional accents: a preliminary investigation. *Journal of Child Language*, **25**: 343-365.

Nelson, K.E. (ed.) (1985). *Children's Language, 5*. Hillsdale, NJ: Lawrence Erlbaum.

Nicol, J.L. (ed.) (2001). *One Mind, Two Languages*. Oxford: Blackwell.

Nicoladis, E. & F. Genesee. (1997). Language development in preschool bilingual children. *Journal of Speech-Language Pathology and Audiology* **21**(4): 258-270.

Nittrouer, S., Studdert-Kennedy, M., & McGowan, R.S. (1989). The emergence of phonetic segments: evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research*, **32**: 120-132.

Nolan, F. (1978). The 'coarticulation resistance' model of articulatory control: solid evidence from English liquids. *Nottingham Linguistic Circular*, **7**(**1**): 28-51.

_______ (1982). The role of action theory in the description of speech production. *Linguistics*, **20**: 287-308.

Nygaard, L.C. Sommers, M.S. & Pisoni, D.B. (1994). Speech perception as a talker-contingent process. *Psychological Science,* **5**: 42-46.

Obler, L.K., Zatorre, R.J., Galloway, L. & Vaid, J. (2000). Cerebral lateralisation in bilinguals: methodological issues. In L. Wei. (ed.). *The Bilingualism Reader.* London & New York: Routledge. pp. 381-393.

Obrecht, D.H. (1968). *Effects of the Second Formant on the Perception of Velarization Consonants in Arabic.* Paris: Mouton.

O'Connor, J.D. (1973). *Phonetics*. Harmondsworth: Pelican.

Ohala, J.J. (1983). The origin of sound patterns in vocal tract constraints. In P.F. MacNeilage (ed.) *The Production of Speech*. New York: Springer-Verlag. pp. 189-216.

_______ (1997). The relation between phonetics and phonology. In W.J. Hardcastle & J. Laver (eds.). *The Handbook of Phonetic Sciences*. Oxford: Blackwell. pp. 674-694.

Oller, D.K. (1975). Simplification as the goal of phonological processes in child speech. *Language Learning*, **24**: 299-303.

Omar, M.K. (1973). *The Acquisition of Egyptian Arabic as a Native Language*. Paris: Mouton.

Padilla, A.M. & Liebman, E. (1975). Language acquisition in the bilingual child. *Bilingual Review* 2: 34-55.

Paradis, J. (1996). Phonological differentiation in a bilingual child: Hildegard revisited. In A. Stringfellow, D. Cahana-Amitay, E. Hughes, & A. Zukowski, (eds.), *BUCLD 20 Proceedings*. Somerville, MA: Cascadilla Press. pp.528–539.

_______ (2001). Do bilingual two-year-olds have separate phonological systems? *International Journal of Bilingualism*, **5**(**1**): 19-38.

_______ & Genesee, F. (1996). Syntactic acquisition in bilingual children: autonomous or interdependent? *Studies in Second Language Acquisition*, **18**: 1-25.

Paradis M. (ed.) (1983). *Readings on Aphasia in Bilinguals and Polyglots*. Montreal: Didier.

_______ (ed.) (1987). *The Assessment of Bilingual Aphasia*. Hillsdale, NJ: Lawrence Erlbaum.

_______ (1997). The cognitive neuropsychology of bilingualism. In A. de Groot & J. Kroll (eds.). *Tutorials in Bilingualism*. Mahwah, NJ: Lawrence Erlbaum. pp. 331-54.

_______ (1998). Language and communication in multilinguals. In B. Stemmer & H.A. Whitaker (eds.). *Handbook of Neurolinguistics*. San Diego: Academic Press. PP. 418-431.

_______ (2000). Language lateralisation in bilinguals: enough already. In L. Wei (ed.). *The Bilingualism Reader*. London & New York: Routledge. pp. 394-401.

Payne, A.C. (1980). Factors controlling the acquisition of the Philadelphia dialect by out-of-state children. In W. Labov, (ed.) *Locating Language in Time and Space*. New York: Academic Press. pp. 143-178.

Pearson, B.Z., Fernandez, S.C. & Oller, D.K. (1993). Lexical development in bilingual infants and toddlers. *Language Learning*, **43**: 93-120.

_______, Navarro, A.M. & Gathercole, V.M. (1995). Assessment of phonetic differences in bilingual-learning infants. *Proceedings of the 19th Annual Boston University Conference on Language Development*. Boston: Cascadilla Press, ii. pp. 427-438.

Perkell, J.S. & Klatt, D.H. (eds.) (1986). *Invariance and Variability in Speech Processes*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Peterson, G.E. & Barney, H.L. (1952). Control methods used in a study of vowels. *Journal of the Acoustical Society of America*, **24**: 175-184.

Petitto, L.A., Katerelos, M., Levy, B.G., Gauna, K., Tétreault, K., & Ferraro, V. (2001). Bilingual signed and spoken language acquisition from birth: implications for the mechanisms underlying language acquisition. *Journal of Child Language*, **28**: 453-496.

Petyt, K.M. (1985). *Dialect and Accent in Industrial West Yorkshire*. Amsterdam: John Benjamins.

Pisoni, D.B. (1997). Some thoughts on normalization in speech perception. In K. Johnson & J.W. Mullennix (eds.) *Talker Variability in Speech Processing*. San Diego: Academic Press. pp. 9-32.

_______ & Lively, S.E. (1995). Variability and invariance in speech perception: a new look at some old problems in perceptual learning. In W. Strange (ed.). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. Baltimore: York Press. pp. 433-462.

Poplack, S. & Meechan, M. (1995). Patterns of language mixture: nominal structure in Wolof-French and Fongbe-French bilingual discourse. In L. Milroy & P. Muysken (eds.), *One Speaker, Two Languages: Cross-disciplinary Perspectives on Codeswitching*. Cambridge: Cambridge University Press. pp. 199-233.

Port, R.F. & Mitleb, F.M. (1983). Segmental features and implementation in acquisition of English by Arabic speakers. *Journal of Phonetics*, **11**: 219-229.

_______ & Rotunno, R. (1979). Relation between voice onset time and vowel duration. *Journal of the Acoustical Society of America*, **66**: 654-622.

Preston, M.S., Yeni-Komshian, G. & Stark, R. (1967). Voicing in initial stop consonants produced by children in the pre-linguistic period from different language communities. Annual Report, 2: 305-323. Neurocommunications Laboratory, Baltimore, MD: Johns Hopkins University School of Medicine.

Pulgram, E. (ed.) (1957). *Studies Presented to Joshua Whatmough*. The Hague: Mouton.

Quay, S. (1995). The bilingual lexicon: implications for studies of language choice. *Journal of Child Language*, **22**: 369-87.

Radwan, M. (1996). *An Experimental Investigation of the Acoustical Temporal Correlates of Voicing Contrast in Stop Consonants (with Reference to Arabic).* Unpublished PhD dissertation. University of Essex.

Rapport, R.L., Tan, C.T. & Whitaker, H.A. (1983). Language function and dysfunction among Chinese and English speaking polyglots: cortical stimulation, wada testing, and clinical studies. *Brain & Language*, **18**: 342-366.

Redlinger, W. E. & Park, T. (1980). Language mixing in young bilinguals. *Journal of Child Language*, **7**: 337-52 .

Reid, E. (1978). *Social and stylistic variation in the speech of children: some evidence from Edinburgh*. In P.J. Trudgill (ed.) *Sociolinguistic Patterns in British English.* London: Arnold. pp. 158-171.

Riney, T.J. & Takagi, N. (1999). Global foreign accent and voice onset time among Japanese EFL speakers. *Language Learning*, **49(2)**: 275-302.

Roach, P. (1991). *English Phonetics & Phonology: A Practical Course. (2nd ed.).* Cambridge: Cambridge University Press.

Roberts, J. (1997). Acquisition of variable rules: a study of (-t, d) deletion in preschool children. *Journal of Child Language*, 24: 351-372.

________ and Labov, W. (1995). Learning to talk Philadelphian: Acquisition of short a by preschool children. *Language Variation and Change*, **7**: 101-12.

Robertson, D.A. & Gernsbacher, M.A. (1998). fMRI exploration of the cognitive processes of mapping. Paper presented at the meeting of the Psychonomic Society, Dallas, Texas.

Rocca, P.D. & Marcelino, M. (1999). Some characteristics of VOT in plosives produced by speakers of English and Portuguese. *Proceedings of the 14th International Congress of the Phonetic Sciences*. University of California, Berkeley. pp. 1425-1428.

Romaine, S. (1984). *The Language of Children and Adolescents: the Acquisition of Communicative Competence*. Oxford: Blackwell.

________ (1995). *Bilingualism*. Oxford: Blackwell.

Ronjat, J. (1913). *Le Dévelopement du Langage Observé chez un Enfant Bilingue*. Paris: Champion.

Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, **274**: 1926-1928.

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

Sander, E.K. (1972). When are speech sounds learned? *Journal of Speech and Hearing Disorders*, **37**, 55-63.

Satz, P. (1979). A test of some *models of hemispheric* speech organisation in the left- and right- handed. *Science*, **203**: 1131-1133.

Schnitzer, M.L. & Krasinski, E. (1994). The development of segmental phonological production in a bilingual child. *Journal of Child Language*, **21**: 585-622.

________ & _______ (1996). The development of segmental phonological production in a bilingual child: A contrasting second case. *Journal of Child Language*, **23**: 547 – 571.

Schreuder R. & Weltens B. (eds.) (1993). *The Bilingual Lexicon*. Amsterdam: John Benjamins.

Scobbie, J.M. (2002) Flexibility in the face of incompatible English VOT systems. Paper presented at the Labphon 8 Conference, Yale University, June 2002.

Scobbie, J.M., Gibbon, F., Hardcastle, W.J. & Fletcher P. (2000). Covert contrast as a stage in the acquisition of phonetics and phonology. In M. Broe and J. Pierrehumbert (eds.) *Papers in Laboratory Phonology V: Language Acquisition and the Lexicon*. Cambridge: Cambridge University Press. pp. 194-207.

_______, Turk, A.E. & Hewlett, N. (1999). Morphemes, phonetics and lexical items: the case of the Scottish Vowel Length Rule. *Proceedings of the XIVth International Congress of Phonetic Sciences*. University of California, Berkeley. pp. 1617-1620.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, **10**: 209: 231.

Shaheen, K. (1979). *The Acoustic Analysis of Arabic Speech*. PhD Thesis. University of Wales.

Sharwood Smith, M. (1991). Language modules and bilingual processing. In E. Bialystok (ed.) *Language processing in bilingual children*. Cambridge: Cambridge University Press, pp. 10-24.

Simon, C. (1976). *A Developmental Study of Acoustic Pattern Production and Perception in Voiced-Voiceless Oppositions*. Unpublished PhD Thesis, University of London.

_______ & Fourcin, A. (1978). Cross-language study of speech-pattern learning. *Journal of the Acoustical Society of America*, **63** (3): 925-935.

Slawinski, E.B. & Wiigs, M. (1999). Categorisation of bilabial stop consonants by bilingual speakers of English and Polish. *Proceedings of the 14th International Congress of the Phonetic Sciences*. University of California, Berkeley. pp. 771-774.

Smith, B.L. (1978a). Effects of place of articulation and vowel environment on 'voiced' stop consonant production. *Glossa*, **12**: 163:175.

_______ (1978b). Temporal aspects of English speech production: a developmental perspective. *Journal of Phonetics*, **6**: 37-67.

_______ & Kenny, M.K. (1999). A longitudinal study of the development of temporal properties of speech production: data from four children. *Phonetica*, **56**: 73-102.

Smith, N.V. (1973). *The Acquisition of Phonology: A Case Study*. Cambridge: Cambridge University Press.

Snow, C.E. & Hoefnagel-Höhle, M. (1977). Age differences in the pronunciation of foreign sounds. *Language & Speech*, **20**: 357-365.

Snow, D. (1997). Children's acquisition of speech timing in English: a comparative study of voice onset time and final syllable vowel lengthening. *Journal of Child Language*, **24(1)**: 35-56.

Sproat, R. & Fujimura, O. (1993). Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics*, **21**: 291-311.

Stager, C.L. & Werker, J.F. (1997). Infants listen for more phonetic detail in speech perception than in word learning tasks. *Nature*, **388**: 381-382.

Stampe, D. (1979). *A Dissertation on Natural Phonology*. New York: Garland.

Stark, R.E. (1980). Stages of speech development in the first year of life. In G. Yeni-Komshian, J.F. Kavanagh, & C.A. Ferguson (eds.). *Child Phonology, 1: Production*. New York: Academic Press. pp. 73-92.

Stemmer B. & Whitaker H.A. (eds.) (1998). *Handbook of Neurolinguistics*. San Diego: Academic Press.

Stevens, K. N. (1998). *Acoustic Phonetics*. Cambridge, MA.: MIT Press.

Stevens, K.N. & Klatt, D.H. (1974). Role of formant transitions in the voiced-voiceless distinction of stops. *Journal of the Acoustical Society of America*, **55**(3): 653-659.

Stoddart, J., Upton, C. & Widdowson, J.D.A. (1999). Sheffield dialect in the 1990s: revisiting the concept of NORMs. In P. Foulkes & G.J. Docherty (eds.) *Urban Voices*. London: Arnold. pp. 72-89.

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 Congress of the Phonetic Sciences*. University of California, Berkeley. pp. 2485-2488.

_______ & Cooper, J.A. (1984). Patterns of lexical and phonological development. *Journal of Child Language*, **11**: 247-71.

Strange W. (ed.) (1995). *Speech Perception and Linguistic Experience: Theoretical and Methodological Issues*. Baltimore: York Press.

_______ (1999). Levels of abstraction in characterising cross-language phonetic similarity. *Proceedings of the 14th International Congress of the Phonetic Sciences.* University of California, Berkeley. pp. 2513-2519.

Stringfellow, A. Cahana-Amitay, D. Hughes, E. & Zukowski, A. (eds.) (1996). *BUCLD 20 Proceedings*. Somerville, MA: Cascadilla Press.

Stuart-Smith, J. (1999). Glasgow: accent and voice quality. In P. Foulkes & G. J. Docherty (eds.). *Urban Voices*. London: Arnold. pp. 203-222.

Studdert-Kennedy, M. (1987). The phoneme as a perceptuo-motor structure. In A. Allport, D. Mackay, W. Prinz, & E. Scheerer (eds.). *Language Perception and Production*. New York: Academic Press.

Swain, M. (1972). *Bilingualism as a First Language*. Unpublished PhD dissertation, University of California, Irvine.

________ (1977). Bilingualism, monolingualism, and code acquisition. In W. Mackey & T. Andersson (eds.). *Bilingualism in Early Childhood*. Rowley, MA: Newbury House.

Thomas, E. (2000). Spectral differences in /ai/ offsets conditioned by voicing of the following consonant. *Journal of Phonetics*, **28**: 1-25.

Tollfree, L. (1999). South East London English: discrete versus continuous modeling of consonantal reduction. In P. Foulkes & G. J. Docherty (eds.). *Urban Voices*. London: Arnold. pp. 163-184.

Treffers-Daller, J. (1998). Variability in code-switching styles: Turkish German code-switching patterns. In R. Jacobson (ed.). *Code-switching Worldwide*. Berlin: Mouton de Gruyter. pp. 177-200.

Trudgill, P.J. (1974). *The Sociolinguistic Differentiation of English in Norwich*. Cambridge: Cambridge University Press.

_________ (ed.) (1978). *Sociolinguistic Patterns in British English*. London: Arnold.

_________ (1986). *Dialects in Contact*. Oxford: Blackwell.

_________ (1999). Norwich: endogenous and exogenous linguistic change. In P. Foulkes & G. J. Docherty (eds.). *Urban Voices*. London: Arnold. pp. 124-140.

Tucker, G.R. (1998). A global perspective on multilingualism and multilingual education. In J. Cenoz & F. Genesee (eds.). *Beyond Bilingualism: Multilingualism and Multilingual Education*. Clevedon: Multilingual Matters. pp. 3-15.

Verma, M.K., Firth, S. & Corrigan, K. (1992). The developing phonological system of Panjabi/Urdu speaking children learning English as a Second Language in Britain. In J. Leather & J. Allan (eds.). *New Sounds. Proceedings of the 1992 Symposium on the Acquisition of Second-Language Speech*. University of Amsterdam. Pp. 174-199.

Vihman, M. M. (1978). Consonant harmony: its scope and function in child language. In J.H. Greenberg (ed.). *Universals of Human Language*. Stanford, CA: Stanford University Press.

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

________ (1992). Early syllables and the construction of phonology. In C.A. Ferguson, L. Menn & C. Stoel-Gammon (eds.). *Phonological Development: Models, Research, Implications*. Timonium, MD: York Press. pp. 439-456.

________ (1996). *Phonological Development: The Origins of Language in the Child*. Oxford: Blackwell.

________ & McCune, L. (1994). When is a word a word? *Journal of Child Language*, **21**: 517-542.

________ & Velleman, S. (2000). Phonetics and the origin of phonology. In N. Burton-Roberts, P. Carr & G.J. Docherty (eds.). *Phonological Knowledge: Conceptual and Empirical Issues*. Oxford: Oxford University Press. pp. 305-339.

________, Ferguson, C.A., & Elbert, M. (1986). Phonological development from babbling to speech: Common tendencies and individual differences. *Applied Psycholinguistics*, **7**: 3-40.

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

von Raffler-Engel, W. & Leburn, Y. (eds.) (1976) *Baby Talk and Infant Speech.* Amsterdam: Swets & Zeitlinger.

Waterson N. & Snow C. (eds.) (1978). *The Development of Communication*. London: Wiley.

Watson, I. (1991). Phonological processing in two languages. In E. Bialystok (ed.) *Language Processing in Bilingual Children*. Cambridge: Cambridge University Press. pp. 25-48.

_______ (1995). The effect of bilingualism on the acquisition of perceptual categories underlying the voicing contrast. *Proceedings of the 13th International Congress of the Phonetic Sciences,* **2**: 710- 713. University of Stockholm.

Watson, J. (1999). Remarks and replies. The directionality of emphasis spread in Arabic. *Linguistic Enquiry*, **30(2)**: 289-300.

Weismer, G. (1979). Sensitivity of voice-onset time (VOT) measures to certain segmental features in speech production. *Journal of Phonetics,* **7**: 197:204.

_______ (1980). Control of the voicing distinction for intervocalic stops and fricatives: some data and theoretical considerations. *Journal of Phonetics,* **8**: 427-438.

Wells, J.C. (1982). *Accents of English* (3 vols.). Cambridge: Cambridge University Press.

Weinreich, U. (1953). Languages in contact: findings and problems. New York: The Linguistic Circle of New York.

_______, Labov, W. & Herzog, M. (1968). Empirical foundations for a theory of language change. In W.P. Lehmann & Y. Malkiel (eds.). *Directions for Historical Linguistics*. Austin, Texas: University of Texas Press. pp. 95-188.

Werker, J.F. (1989). Becoming a native speaker. *American Scientist,* **77**: 54-59.

_______(1995). Age-related changes in cross-language speech perception: standing at the crossroads. In W. Strange (ed.). *Speech Perception and Linguistic Experience: Issues in Cross-linguistic Research.* Baltimore: York Press. pp. 155-170.

_______ & Tees, R.C. (1984). Cross-language speech perception: evidence for perceptual reorganization during the first year of life. *Infant Behaviour and Development,* **7**: 49-63.

_______ & Logan, J.S. (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics*, **37**: 35-44.

_______, Gilbert, J.H.V., Humphrey, K. & Tees, R.C. (1981). Developmental aspects of cross-language speech perception. *Journal of the Acoustic Society of America*, **75**: 1866-78.

White, L. & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in second language acquisition. *Second Language Research*, **12**: 233-265.

Whitworth, N. (2000). Acquisition of VOT and vowel length by English-German bilinguals: a pilot study. *Leeds Working Papers in Linguistics and Phonetics,* **8**: 229-243.

_______ (forthcoming). Speech rhythm production in three German-English bilingual families. To appear in *Leeds Working Papers in Linguistics and Phonetics,* **9.**

Williams, A. & P. Kerswill (1999). Dialect levelling: continuity vs. change in Milton Keynes, Reading and Hull. In P. Foulkes & G. J. Docherty (eds.). *Urban Voices.* London: Arnold. pp. 141-162.

Windsor, F., Kelly, M.L. & Hewlett, N. (eds.) (2002). *Investigations in Clinical Phonetics and Linguistics*. Mahwah, NJ.: Lawrence Erlbaum.

Windsor Lewis, J. (ed.) (1995). *Studies in General and English Phonetics in Honour of Professor J D O'Connor* London: Routledge Kegan Paul.

Wode, H. (1997). Perception and production in learning to talk. In M. Young-Scholten & S.J. Hannahs (eds.), *Focus on Phonological Acquisition*. Amsterdam: John Benjamins. pp. 17-46.

Yavas, M. (2002). Voice onset time patterns in bilingual phonological development. In F. Windsor, M.L. Kelly & N. Hewlett (eds.). *Investigations in Clinical Phonetics and Linguistics*. Mahwah: Laurence Erlbaum Associates Inc. pp. 341-350.

Yeni-Komshian, G.H., Caramazza, A. & Preston, M.S. (1977). A study of voicing in Lebanese Arabic. *Journal of Phonetics*, 5: 35-48.

________ Kavanagh, J.F. & Ferguson, C.A. (eds.) (1980). *Child Phonology. Volume 1: Production*. New York: Academic Press.

Younes, M. (1993). Emphasis spread in two Arabic dialects. In Eid, M. & Holes, C. *Perspectives on Arabic Linguistics V*. Amsterdam: John Benjamins. pp. 119-145.

Young-Scholten, M. & Hannahs, S.J. (eds.) (1997). *Focus on Phonological Acquisition*. Amsterdam: John Benjamins.

Zampini, M.L. & Green, K.P. (2001). The voicing contrast in English and Spanish: the relationship between perception and production. In J.L. Nicol (ed.) *One Mind, Two Languages*. Oxford: Blackwell. pp. 23-48.

Zlatin, M. & Koenigsknecht, R. (1976). Development of the voicing contrast: A comparison of voice onset time values in stop perception and production. *Journal of Speech and Hearing Research*, 19: 92-111.