

Codeswitching in Mirpuri speaking Pakistani heritage preschool children: Bilingual language acquisition

Sean Pert,¹ Carolyn Letts²

¹ Rochdale Primary Care Trust, Rochdale Sure Start Project
and University of Newcastle, U.K.

² University of Newcastle, U.K.

Acknowledgments

Data collection for this project was aided by a grant from the National Special Interest group in Bilingualism, Royal College of Speech and Language Therapists, U.K. and the Rochdale Sure Start projects. The study would have been impossible without the support and hard work of, Shazye Iqbal, Nazmeen Kausar, Safina Kouser, Tahira Mahmood, and Zahida Warriach, Bilingual Speech and Language Therapy Assistants, in collecting, translating and discussing the data. Thanks go to and Mark Sebba and Paul Robinson for assistance in the use of the CLAN tools and LIDES coding. We also wish to thank the staff, parents and children from the nurseries, schools and clinics in Rochdale for allowing us to collect the data. Finally we are very grateful to Professor Carol Myers-Scotton and to Dr. Ghada Khattab for comments on earlier drafts of this article.

Abstract

In an environment where bilingual adults codeswitch routinely, children are likely to show codeswitching in their own utterances from an early age. However, child language has rarely been studied from this viewpoint and codeswitching may be viewed negatively by many educators and other professionals involved with bilingual children. This research examines codeswitching recorded during the assessment of two groups of preschool Mirpuri speaking children living in Rochdale, U.K. The children's codeswitched utterances are examined using the Myers-Scotton Matrix Language Frame (MLF) and 4M models. The results show that codeswitching was common. Codeswitched utterances were generally longer than utterances that did not display codeswitching and they conformed to the MLF/4M models. Codeswitched utterances may be the most representative utterances of a bilingual child's expressive potential; only when speaking to a fellow bilingual who accepts codeswitching is the bilingual child able to express him- or herself to this potential. Codeswitched utterances, and not solely L1 or L2 utterances, should therefore form the basis for establishing bilingual norms of bilingual language acquisition for populations where codeswitching is effectively the L1.

Key words

bilingualism
codeswitching
language acquisition

Address for correspondence

Carolyn Letts, School of Education, Communication and Languages Science, University of Newcastle, King George VI Building, Queen Victoria Rd, Newcastle upon Tyne, NE1 7RU, U.K.; e-mail: <c.a.letts@ncl.ac.uk>.

1 Introduction

Codeswitching is a phenomenon of language contact. When speakers have access to more than one language code they have the potential to alternate the language codes or to use both codes together. The manner and degree in which this occurs in reality has been the subject of much research interest. Codeswitching literature highlights that codeswitching may be viewed from two main perspectives: the social discourse functions that codeswitching signals and the grammatical outcomes of codeswitching (Bentahila & Davies, 1995; Boumans, 2001; Boztepe, 2003). In a review, Auer (1998) concludes firstly, that speech communities indicate group membership via the use of codeswitching and secondly, that intrasentential codeswitching is syntactically and morphologically constrained. This contrasts with historical and popular views of codeswitching, where the use of two languages in the same communicative episode has often been seen as a problem, with claims by some professionals that codeswitching signifies deficient language skills, poor language differentiation and even a sign of deviation from the bilingual norm (Brice & Anderson, 1999; Heredia & Altarriba, 2001). A recent review by Chiocca (1998) indicated that even young preschool bilingual children were able to differentiate languages and utilize codeswitching in a pragmatically sophisticated manner. In fact, numerous studies have demonstrated that codeswitching between bilingual speakers is a natural, common and useful communication strategy (Reyes, 1995). Codeswitching is however largely discussed as a phenomenon of adult language. Where it appears in the language of young bilingual children, it is often assumed to be a form of language *mixing*, something which disappears as the child gains competence in both languages (Jisa, 2000; Redlinger & Park, 1980 and see also the discussion in Lanza, 1997). It should be noticed that these studies are concerned with simultaneous bilingual acquisition, where the child hears clearly separated input from two or more languages from an early age. In the case of a child who is acquiring language in a community where codeswitching is the norm, or where a codeswitched variety is the child's first language, it will be more difficult to distinguish early mixing from developing codeswitching.

The term "codeswitching" may be used in several different ways. This may lead to confusion as similar or identical terms are used for a range of communication behaviors. Li Wei recognizes that codeswitching is a continuum: "A long narrative may be divided into different parts which are expressed in different languages; sentences may begin in one language and finish in another; words and phrases from different languages may succeed each other" (2000, p. 16). For some authors the term is not associated solely with bilingualism. Leonard uses the term to describe how monolingual children "...modify their speech style and complexity as a function of the listener's age, language ability and social status" (1998, p. 82).

Authors that employ differing terminology include Cunningham-Andersson and Andersson (2004, p. 152) who label "changing from one language to another" as "language switching" in contrast to "language mixing" which they describe as "using words from more than one language in a single utterance." "Codeswitching" has also been used to describe language mixing that reflects pragmatic competence in the bilingual speaker (Cheng, 2003; Isaac, 2002).

The terminology most often reflects the use of two (or more) languages within the same utterance (Abudarham, 1987; Heredia & Altarriba, 2001; Myers-Scotton, 2002). Specific use of “intersentential” as opposed to “intrasentential” codeswitching avoids confusion. *Intersentential codeswitching* is when a speaker produces an utterance in one language, and then the subsequent utterance in another language. *Intrasentential codeswitching* occurs when lexical items or grammatical morphemes from two or more language codes are incorporated into the same sentence, or more specifically, the same clause (Paradis, Nicoladis, & Genesee, 2000; Shin, 2002). In this paper, we will use the term ‘codeswitching’ to refer only to intrasentential codeswitching. This is in common with the definition proposed by Bhatt, who states that intrasentential codeswitching is “...the alternate use of two linguistic systems within a clause” (Bhatt, 1997, p. 223). Strictly speaking, we should refer to *intraclause* codeswitching, as the focus is on switching within the clause and does not include switching that occurs at clause boundaries in complex sentences (see Myers-Scotton, 2002). The term intrasentential codeswitching will be used here however, as this appears to be commonly used by a range of researchers. Moreover, the developmental data described here only contain single clause utterances.

Myers-Scotton has proposed that codeswitching should be viewed as a construction where “...the Matrix Language is the source of structure for the grammatical frame of the bilingual clause. The other language, the Embedded Language, may supply content morphemes that are inserted in this frame.” (2002, pp. 19–20). The Matrix Language Framework model (MLF) is a comprehensive model of bilingual language competence and production. It recognizes that codeswitching is a highly constrained process and not random or impaired language production. Myers-Scotton states that the model demonstrates that codeswitching can be shown to have a “...*permissible* grammatical structure and how such constraints relate to general linguistic structure” (2002, p. 11, *our italics*). Researchers have already demonstrated that children’s codeswitching obeys the MLF model “...the majority of the time” for French-English bilingual children (Paradis, Nicoladis, & Genesee, 2000, p. 245. Note that in some respects Paradis et al. apply different criteria from Myers-Scotton in their identification of codeswitching; however their data support the constraints predicted by the MLF model). Further study of children using other language combinations is important in order to gain insights into the typical acquisition of codeswitching and codeswitched varieties.

Another aspect associated with the use of different linguistic codes is the overlap between codeswitching and lexical borrowing. Borrowing is a well-established phenomenon in language contact situations, where one language will readily absorb lexical items from another. Where borrowed words represent concepts or entities for which the “borrowing” language does not have a word, then they can be clearly recognized as such. The problem arises however where lexical equivalents appear to exist in both languages, but one form (the “borrowed” form) may have become the preferred option for the community. It has been claimed that borrowings may be differentiated from codeswitching by the virtue of the level of integration into the base language. Poplack (1980) proposed levels of integration of items in terms of the phonological, morphological and syntactic integration of items. However, in practice this has proved problematic. For example, Romaine states: “in the case of Panjabi/English mixed compounds, a great

many are not phonologically integrated, and they differ in the extent to which they are morphologically and syntactically integrated” (Romaine, 1995, p. 144). Frequency data on an item may give a clearer indication as to whether a lexical item from one language has supplanted another; one example is the work reported by Shin (2002), where frequency data is used to determine the status of single noun switches in the context of an utterance.

The differentiation of codeswitching and borrowing has theoretical implications, as counterexamples to various proposed grammatical constraints and matrix-embedded language models have been dismissed as examples of borrowing (Sankoff, 1998). However, Myers-Scotton (2002) argues that the distinction is irrelevant, as lexical borrowings are subject to the same morphosyntactic constraints as codeswitched items. A further implication however is that increasing preference for a lexical item may be a reflection of language shift in the speaker, such that eventually the item is only readily available to them in this “borrowed” form. Martin, Krishnamurthy, Bhardwaj and Charles (2003) document data indicative of language change in children from the Panjabi-speaking population in the West Midlands, U.K. Lexical and grammatical features suggest that these changes are the result of the influence of English. Evidence for such shift will be important in trying to track typical language acquisition patterns in bilingual children; the loss and replacement of first language (L1) items may be the norm and reflected as such in the codeswitching patterns shown. Here we look at patterns of lexical usage as an indicator of shift in preferred language; this is not an attempt to distinguish codeswitched from borrowed items.

1.1

Bilingual language development

Nicoladis and Genesee point out that “...most research has focused on largely middle-class families in which one parent usually speaks one language with the child while the other parent usually speaks another language — the so-called one parent-one language context” (1997, p. 259). This approach has dominated the thinking on bilingual language acquisition, despite the fact that it is not typical of bilingual language acquisition (*ibid*). Even if this approach were common, it has been shown that the arrangement is an illusion, as parents’ behavior does not always conform to their own expectations. Goodz (1989, p. 32) found that “...all the parents spoke both languages to their children under some conditions. This was true even in parents describing themselves as firmly committed to maintaining a clear parent-by-language distinction.” It is therefore extremely difficult to control for the relative language exposure children receive, even on a small scale. It is also questionable that results from such studies are applicable to the many children who become bilingual through different routes and in different contexts.

A number of researchers have suggested that there are grammatical constraints operating on codeswitching (see e.g. Azuma, 1993; Myers-Scotton, 2002; Pfaff, 1979; Poplack, 1980). This study aims to look at such constraints using Myers-Scotton’s Matrix Frame Language (MLF) and 4-M models (*ibid*) and applying them to the naturalistic bilingual language development of a group of children in Rochdale, U.K. (see 2.1 for further details of this community). These children are thought to be developing from predominantly monolingual Mirpuri speakers to bilingual Mirpuri-English speakers.

However, unlike previous one-person-one-language approaches, the study examines bilingual acquisition in an uncontrolled situation. This raises some important questions about the way codeswitching is conceptualized. If children acquire language from codeswitching adults, then the children will not have a comparable L1 acquisition to their peers in Pakistan. There are no systematic studies available, but it appears that adults in Rochdale U.K. frequently use intrasentential codeswitching with an L1 Mirpuri word order and grammar. This contrasts with other bilingual acquisition situations where L1 and L2 are viewed to be completely separate and distinct. As a result of this, English lexical items will not necessarily be sociolinguistically marked as L2 items. These children may therefore be acquiring codeswitched Mirpuri-English as their first language.

Preschool children in Rochdale generally have only minimal contact with non-codeswitched English prior to reception class entry. It would be useful to discover how quickly (if at all) intrasentential codeswitching becomes established in the expression of these children. If intrasentential codeswitching does occur in young children and the MLF/4M models are valid as universal models of codeswitched language, then it would be expected that the pattern for Mirpuri-English codeswitching is rule bound in the same way as codeswitching for any other language pair. If this is not the case, and the language combinations lead to different outcomes, then the MLF/4M models would not be as universal as is claimed.

The characteristics of codeswitching in this population form an interesting case study as the population are not elite bilinguals. Observation suggests that when codeswitching is employed by children it is not rejected by adults. It is not clear how any transfer of language dominance is viewed by parents. Anecdotal evidence from speech-language therapists working with the community suggests that outcomes of bilingualism or loss of L1 are both common and accepted.

If the MLF model applies to young Mirpuri speakers' early grammar, it is predicted that Mirpuri will form the matrix language, since contact with non-codeswitched English has been minimal up to nursery school entry. If this is the case, word order should therefore reflect Mirpuri word order (SOV) and content morphemes may be drawn from either Mirpuri or English. Late system morphemes as defined by the 4M model, should always be drawn from the matrix language (Mirpuri). Late system morphemes "...depend on their form on information *outside* their immediate maximal projection" (Myers-Scotton, 2002, p. 75). These include subject-verb (gender) agreement morphemes.

As the status of codeswitched utterances in relation to the child's overall language acquisition is not clear, this study compares mean length of utterance (MLU) for assumed L1 (Mirpuri), L2 (English) and codeswitched utterances. Such comparison will give some insight as to whether the children's true L1 is actually codeswitched Mirpuri-English, as used by the adults in their community. In addition, in order to explore the status of codeswitched items in terms of language shift, the distribution patterns of lexical items from the two languages will be discussed. According to the MLF model, content morphemes would be expected to range from use only in one language to equal use. Similarly, the 4M model indicates that late system morphemes should be polarized

and originate from the matrix-language only. A high proportion of English content morphemes may mark the beginning of language shift for the individual.

The study aimed then to explore the following questions:

1. Does Mirpuri-English codeswitching occur in the language of preschoolers with a Mirpuri-speaking home background?
2. If intrasentential codeswitching does occur, is this subject to grammatical constraints, such that it conforms, even with young children, to the MLF/4M models?
3. In terms of frequency of English items embedded into utterances otherwise spoken in Mirpuri, is there any evidence of a wholesale language shift towards English?
4. What are the implications for norms of bilingual language acquisition, and for language assessment?

In addition, data were subdivided into two age groups to ascertain whether there are any apparent changes over time as the child's language continues to develop.

2 Method

2.1

The Pakistani heritage population as a case study

7.9% of the U.K. population describe themselves as belonging to a minority ethnic group. The Pakistani heritage population make up 1.3% of the population of the U.K. or 16.1% of the minority ethnic population (National Statistics, 2004a). The Pakistani and Bangladeshi heritage populations, make up the most widely dispersed ethnic minority groups throughout England (National Statistics, 2004b). However, ethnicity does not always equate with bilingualism; there are currently no data on languages spoken in U.K. households as the national census does not request this information. Pakistani heritage children aged 0–14 years make up 34% of that population compared with only 19% for the population of white children (Scott, Pearce, & Goldblatt, 2001, p. 11). This means that the Pakistani heritage population, like the rest of the minority ethnic population, is demographically much younger and experiencing a higher birth rate than the white population.

Rochdale is situated in the North West of England, north east of Manchester. 7.7% of the Rochdale population is from a Pakistani Heritage background (National Statistics, 2003). One in five nursery and primary school aged children in Rochdale are reported to be bilingual, of which approximately 80% are Pakistani heritage¹.

2.2

The linguistic context

The linguistic background of this population is complex. Three Pakistani heritage languages have been identified: Mirpuri, Punjabi, and Urdu (see Pert & Stow, 2001). Of these, Urdu is a high status language both in the U.K. and in the Indian subcontinent

¹ Personal communication, I. Jones, 2004

(described by Romaine (1995, p.307) as having a “long literary tradition”). Punjabi is very close to the Panjabi spoken by Sikhs in the West Midlands U.K., and studied by Martin, Krishnamurthy, Bhardwaj, and Charles (2003: see also the authors’ clarification note, 2004) and by Romaine (1986, 1995). The description *Punjabi* is used for the language of Muslim speakers, and the two use different scripts. While they are considered to be the same language in Pakistan, it is possible that the two have diverged to some extent in the U.K., being spoken by two groups who differ in terms of both religion and location. Mirpuri is not well documented. The label ‘Mirpuri’ is derived from the Pakistani city of Mirpur and its surrounding district. The area is mainly rural, with some industrialization. Many Pakistani heritage populations in the U.K. have origins in the Mirpur District, including those living in Rochdale, Nottingham (Nerwal, 2002) and Sheffield (Reynolds, 2002). The *Linguistic Minorities Project* (1985) mentions that a large number of respondents in their survey “used Mirpuri or another regional variety, but relatively rarely named it as their first language” (p. 143).

It is also unclear whether Mirpuri should be viewed as a dialect (of either Punjabi or Urdu) or as a distinct language. We have been unable to find any systematic studies that would provide an answer to this question. An item posted on a *Linguist List* discussion group in 1999 (Smithers), describes Mirpuri vocabulary as overlapping with both Punjabi and Urdu, but that much of the vocabulary and grammar “have no parallel in Urdu and Punjabi.” In Rochdale the lexical overlap between the three languages has been observed, as have differences in morphology and syntax. As a result, the young children in this study have limited or no understanding of the other two languages and at least some adults have similar difficulties. Mirpuri is considered significantly different enough to be differentiated by the Driving Standards Agency (2005), which regulates driver licensing and testing in the U.K. (In order to gain a driving license, applicants must pass both a practical and theory test, and the tests must be made available in the applicant’s first language. A Mirpuri version of the theory test has been formulated for this purpose.) The British Broadcasting Corporation (BBC) also transmits radio programs in all three Pakistani heritage languages including Mirpuri (British Broadcasting Corporation, 2004a).

The issue is further clouded by socioeconomic factors. Mirpuri is a preliterate language, both in the U.K. and in Pakistan. The lack of a written form linked with the relatively modest socioeconomic background of the speakers is reflected in perceptions of the status of the language. The Mirpuri and Kashmiri community are often viewed negatively (British Broadcasting Corporation, 2004b). Mirpuri speakers frequently report themselves as Urdu speakers (see Pert & Letts, 2003). This may be reinforced by English speakers who are only aware of Urdu and assume this is the language of all Pakistani-origin speakers, since it is the official language of Pakistan. Urdu is “...a status symbol among ‘educated’ classes and those aspiring to upward mobility and is also the official and formal language of the state and of the state education system” in Pakistan (Kirklees Metropolitan Council Education Service and Dewsbury College, 2002, p. 17). Urdu speakers may view Mirpuri as nonstandard and inferior and describe Mirpuri as a “dialect” or “slang” in a pejorative sense.

It is tempting to speculate that the three varieties may be subject to selection by the speaker in accordance with sociolinguistic domain (as described by Romaine, 1995).

This may be the case with the most highly educated speakers in Rochdale, who might use Urdu for more formal activities. However it seems that a sizeable proportion of the population do not have access to all three varieties. The experience of the first author, working as a speech-language therapist in Rochdale, is that adult Mirpuri speakers do not fully understand Punjabi or Urdu speakers and vice versa. One of the bilingual speech language therapy assistants (STLAs: see below) who works with the first author has a good knowledge of the three languages, but chose to learn Urdu at college as an adult.

Grammatical descriptions of Mirpuri do not exist. The first author has been able to derive a grammatical parsing for the stimuli sentences used in this study by two main processes. Firstly, grammars are available for Urdu and for (Sikh) Panjabi (Schmidt, 2004; Bhardwaj, 1995 respectively), and the languages are sufficiently similar to make analogies from one to the other. Secondly, bilingual STLAs who work closely with the first author have acted as informants. The first author has also acquired informally a degree of working knowledge of the languages through his work as a speech and language therapist. Differences between the three varieties are illustrated by the following example. The first is of one of the stimuli sentences used in this study²:

- (1) Target sentence: “the (girl) baby is drinking milk”

MIRPURI:	<i>kuril beibil/qaqa dud pini</i>	<i>pi</i>
ENGLISH:	girl/ baby	milk drinking+FEMALE is+FEMALE
PUNJABI:	<i>kuril beibil qaqa dud pindi</i>	<i>eh</i>
ENGLISH:	girl/ baby	milk drinking+FEMALE is
URDU:	<i>lurkil beibil qaqa dud pi ruhi</i>	<i>he</i>
ENGLISH:	girl/ baby	milk drink doing+FEMALE is

(NB *beibi* or *qaqa* are acceptable variants for “baby” in each of the languages, as spoken in Rochdale.)

Example 2 gives one of the comprehension items in the adapted Derbyshire Language Scheme (DLS) assessment (see below for description of DLS):

- (2) Instruction: “show me the key”

MIRPURI:	<i>miki das jabi</i>
PUNJABI:	<i>menu das jabi</i>
URDU:	<i>muje dikau jabi</i>
ENGLISH:	me show key

The examples show that all three varieties have an SOV word order for declarative sentences, and that there is gender agreement between subject and verb. This may be expressed as an inflection on the lexical verb, plus in Mirpuri the selection of the auxiliary, (see the Mirpuri and Punjabi examples of *drink*). A further verb phrase construction

² The Mirpuri language does not have a written form. Examples in this paper are given using Roman script that constitutes a rough approximation to the forms used. Details of phonetic characteristics can be obtained from the first author.

exists, that of lexical verbal element plus operator plus auxiliary, with gender agreement on the operator and auxiliary. This is illustrated in the Urdu version of *drink*, and is a common construction in South Asian languages (described as “*compound verbs*” by Romaine, 1995, p. 131). Vocabulary differences between the varieties exist for nouns and lexical verbs, and also for pronouns and auxiliaries.

It is important to note that in adults codeswitching has been observed not only between one Pakistani Heritage language and English, but also between two or more Pakistani heritage languages. An utterance may therefore include words and morphemes that are all Pakistani heritage items but drawn from both Mirpuri and Urdu, for example. This further clouds perceptions of which language is being spoken in the U.K. situation, where Mirpuri, Punjabi and Urdu speakers are in language contact situations. However, the children in this study only rarely displayed codeswitching that was not Mirpuri-English in the data collected, and the study will therefore focus on these two languages.

For the purposes of this study, child participants have all been identified as being most comfortable using Mirpuri or a codeswitched variety of Mirpuri and English. Some of the main features of Mirpuri will now be described.

2.3

Mirpuri—basic syntax and grammar

Mirpuri is a *Subject + Object + Verb* (SOV) language (although word order may be less rigid than in English). Male and female gender agreement applies within a simple sentence: the verb phrase looks to the subject noun phrase for the form of the present progressive suffix and the auxiliary verb form (female *i* and male *a*).

Examples

- (3) Female gender marking in Mirpuri:

Kuri kela kha -ni pi
 Girl banana eat -ing + FEMALE is + FEMALE
 ‘(the) girl is eating (a) banana’

- (4) Male gender marking in Mirpuri

Mura kela kha -na pija
 Boy banana eat -ing + MALE is + MALE
 ‘(the) boy is eating (a) banana’

Note that although the examples here have natural gender, grammatical gender applies to all nouns.

Many actions that would be expressed as simple verbs in English are expressed through a compound verb form consisting of a noun-like item and an operator, as described above. In Mirpuri, there are two main operators that carry subject-verb gender agreement in the same way as lexical verbs. They also function as lexical verbs when they occur in isolation, referring to a generic action. These verbs are *kar* and *mar*.

Kar functions (1) as a lexical verb expressing activity, and (2) to carry subject-verb gender agreement in compound verbs. For example, female gender with generic action verb in Mirpuri:

- (5) *Jenani ishara kar -ni pi*
 Lady point DO -ing+FEMALE is+FEMALE
 ‘(the) lady is pointing’

Mar functions (1) as the verb ‘to hit’, (2) as a verb expressing activity, (3) to carry subject-verb gender agreement and (4) to express the concept of contact and is used with items such as ‘brush’ and ‘clap’ where items are brought together. This is illustrated in the following examples.

- (6) *Mar* as the verb ‘to hit’ in Mirpuri:
Jenani mura qi mar -ni pi
 Lady boy to hit -ing+FEMALE is+FEMALE
 ‘(the) lady is hitting (the) boy’

- (7) Female gender with generic action word+contact in Mirpuri:
Jenani kungi mar -ni pi
 Lady comb DO+CONTACT -ing+FEMALE is+FEMALE
 ‘(the) lady is combing’

2.4

The MLF and 4M models:

Predicting normal codeswitching behavior

The model of analysis used is based on the MLF and 4M models. These examine codeswitching in terms of two key oppositions. The first opposition is that of the Matrix Language—Embedded Language. The Matrix Language provides the abstract structure of the utterance. The second opposition is concerned with content morphemes and system morphemes. Content morphemes convey semantic and pragmatic relationships whilst system morphemes are concerned with how the content morphemes relate to each other.

In more recent work (Myers-Scotton, 2002) a complementary model is proposed that divides system morphemes into three types. Hence there are four morpheme types in the utterance (4M):

- content morphemes which receive or assign thematic roles, for example, nouns tend to receive thematic roles (agents, patients and themes), and verbs tend to assign thematic roles.
- three types of system morphemes
 - ◆ early system morphemes which are related to the head of the content morpheme phrase, for example, determiners and plurals.
 - ◆ late system morphemes
 - bridges which create well-formed utterances, for example, ‘of’, as in “piece of cake.”

- outsiders which look outside their immediate environment for their form, for example, subject-verb agreement, gender agreement on adjectives.

The model labels morphemes according to their role and position in the formulation of the utterance. The MLF and 4M models together allow us to predict the pattern of codeswitching in bilingual children.

It is predicted that content morphemes, typically nouns and verbs will constitute the vast majority of codeswitched items. Nouns and verbs are described as "...prototypical content morphemes" (Myers-Scotton, 2002, p. 75). These will be the heads of noun- and verb phrases that either assign or take thematic roles in the sentence. Crucially, while codeswitching of content morphemes may be widespread, sentence structure in terms of word order will be that of the matrix language. As Mirpuri and English diverge in terms of basic sentence structure (SOV vs. SVO), it should be possible to identify which is the matrix language for any particular utterance.

Early system morphemes are dependent on the head of a phrase for their form and include "...phi-features of person, number, and gender in relevant languages." (Myers-Scotton, 2002, p. 75). They will be drawn from the same language as the head. Late system morphemes are of two types, *bridge* and *outsiders*. Bridge morphemes link items to produce well formed utterances and are likely to be in the matrix language. Outsiders look outside the phrase for agreement and hence information on their form. They are likely to be drawn from the matrix language. In Mirpuri these may include the present progressive verb form and auxiliary verb BE (which must both agree with the subject of the sentence). Gender agreement on outsider adjectives should also come from the matrix language; however as an artifact of the elicitation procedure (and the developmentally early stage of the children's language), examples of such adjectives do not appear in the data reported here. Note also that in Mirpuri, there are no determiners.

2.5

Participants

Twenty-five typically developing children were recruited from local nurseries, playgroups and primary schools in Rochdale, U.K. Information on Rochdale and the Pakistani heritage population living there has been published in a previous paper (see Pert & Letts, 2003). Most children live in a designated "Sure Start" area: that is to say, government funding has been allocated to support initiatives dedicated to improving the health and development of young children in the area, who are considered to be at risk from social deprivation (Glass, 2001). Oral and written consent to participate was gained from parents with the assistance of bilingual STLAs. These assistants are employed to work alongside speech and language therapists (SLTs) in the Rochdale area to assist in assessment and intervention with children who do not have English as their first language, and to facilitate communication with families. They have been specially trained in this role and care is taken that there are STLAs available who speak the relevant languages and who can identify which language the child is using. For the purpose of this study, each child's language was established as Mirpuri through informal discussions between the child and STLA prior to the commencement of formal assessment.

Data were collected as part of a larger study to produce normative data for a Pakistani Heritage expressive language assessment. Of the 25 children, 17 produced an expressive language sample and also fell into the age bands 3;06–3;11, (henceforth Group A), or 4;00–4;05 (henceforth Group B). This is the age range during which children are first typically exposed to English in nursery and it is the data from these children that are detailed in this paper. The division into two age bands is an arbitrary measure to establish whether change in language patterns over time is readily identifiable. Table 1 gives details of age and gender.

Table 1

Children included in the study

Group	Age	Mean	Female	Male	Total
A	3;6 – 3;11	3;9	3	3	6
B	4;0 – 4;5	4;2	6	5	11
			9	8	17

2.6

Language sampling

Children’s verbal comprehension was assessed using the Derbyshire Language Scheme Rapid Screening Test—Mirpuri Adaptation (Pert, 2000). This informal procedure was originally developed by Masidlover and Knowles (1979) and was developed to be used alongside a language development program for children with language delay, the *Derbyshire Language Scheme (DLS)*. It assesses children’s ability to comprehend a number of *information carrying words* (essentially content words) in one utterance, and as such is relatively easy to adapt to different languages. While not standardized it gives an idea as to whether an individual child has a level of spoken language understanding similar to his or her age-peers.

Children’s expression was sampled using “*Je Zindegi*” (English ‘*This Life*’), an expressive language assessment for early sentences devised by the first author. The assessment comprises four practice items and 21 stimulus items. Each item is a picture of a child or adult from the Pakistani heritage community carrying out an everyday action, for example, a lady cooking a meal. The assessment development has been described in previous papers (see Pert & Letts, 2003; Pert & Stow, 2001). The child is asked to describe what is happening in the picture. It should be noted that this format targets a core number of sentence structures and is not a measure of spontaneous language. It has the advantage of prompting the child to attempt a set of target structures, and also that the potential target for each utterance is comparatively transparent.

Children’s utterances were transcribed at the time of assessment by trained bilingual STLAs. Each STLA had Mirpuri as a mother tongue. As Mirpuri has no written form and STLAs are not trained in the use of the International Phonetic Alphabet (IPA) (The International Phonetic Association, 1996), it was decided to use Roman script. This had several advantages. It was relatively quick to transcribe children in situ and the scripts could be easily discussed with the researcher. No assumptions were made

about “correct” target forms. The aim was to ascertain how the children would choose to construct a sentence to describe each picture, whether using Mirpuri, English or a codeswitched version.

The assessment took on average 15 mins with each child. Attempts were made to audio-record as many children as possible but some children refused to speak in the presence of a microphone. Those children’s utterances were transcribed at the time of assessment. Other children were recorded using a MDP 500 professional minidisk recorder with a Shure SM58 unidirectional dynamic microphone, in addition to on-the-spot transcription.

2.7

Data analysis

All scripts were transcribed and coded electronically for further analysis. The CLAN (computerized language analysis) program (Spektor & Tuthill, 2000) was used to produce standardized scripts. These were checked for consistency using the CLAN CHECK program and reference was made to both the CLAN manual (MacWhinney, 2003a) and the CHAT manual (MacWhinney, 2003b).

The CLAN scripts were then printed off and rechecked for accuracy of translation by the bilingual speech and language therapy assistants and speech and language therapist. Reference to the audio recording, if available, was made. Any alterations and clarifications were made at this stage (see LIPPS group (Language Interaction in Plurilingual & Plurialectal Speakers), 2000).

For the purposes of this study, all utterances were removed that were ambiguous, incomplete or where the language categorization (L1, L2 or codeswitched) was not clear. Utterances were then categorized into L1 only (utterances that were identifiable as only containing Mirpuri words and morphemes), L2 only (utterances identifiable as containing only English), and intrasentential codeswitched utterances (CS). Only utterances that clearly had at least one word from each language were retained as instances of codeswitching. In order to make meaningful comparisons of MLU, all single word utterances were excluded from MLU calculations. For the younger Group (A), 24/111 (21.6%) of utterances were excluded because they were incomplete, ambiguous or consisted of only a single word. For the older Group (B), 45/254 (17.7%) of utterances were removed.

Frequency of items was also calculated for each group, giving total number of words, total number of different words and the relative frequency of each word. Allomorphs of one word were counted together, for example, “*khalt-a*” (stand-ing + male) and “*khalt-i*” (stand-ing + female). Frequency analysis was used to look for evidence of language shift.

Qualitatively, data were analyzed in terms of grammatical and syntactic structure, considering the Myers-Scotton 4M and Matrix Language Frame models. No attempt was made to differentiate “borrowed” lexical items from codeswitched lexical items. Where there was a lexical item from L2 in an L1 frame the utterance was considered to contain a codeswitch. Utterances containing codeswitching were then categorized in terms of thematic role. It is important to note that the utterances were elicited in a test situation

and were not spontaneous. The distribution of thematic roles is therefore determined to a large extent by the stimulus pictures presented to the children. Codeswitched items were categorized into the thematic roles of agent, patient, verb, location and goal. Theme and patient were not differentiated (see Saeed, 2003 for discussion of these terms).

3 Results

3.1

Verbal comprehension (Mirpuri)

All the children demonstrated satisfactory verbal comprehension abilities in Mirpuri, as assessed using the Mirpuri adaptation of the DLS screening test.

3.2

Expression

Table 2 gives the mean and range for total number of utterances per child, numbers of children displaying intrasentential codeswitching and the mean and range of codeswitched utterances per child, for each of the two groups. It is clear that all the children codeswitched at some point. However, there appears to be little change across the two age groups in terms of the frequency of codeswitching.

Table 2

Number of utterances and intrasentential codeswitched utterances produced

Group	Mean total utterances per child (range)	Children using intrasentential code switching	Mean number of utterances per child containing intrasentential code switching (range)
A (n=6)	27.33 (4 to 39)	6	6.2 (2 to 15)
B (n=11)	31.27 (21 to 49)	11	8.27 (2 to 16)

This indicates that codeswitching is an established and stable pattern of behavior in these children by around three and a half years.

Table 3 shows the distribution of Mirpuri and English words for the total corpus of utterances prior to removing single word and ambiguous utterances. Approximately a quarter of types and tokens were identifiable as English words.

Table 4 shows the proportions of multiword utterances that were L1 only, L2 only and codeswitched (CS) for Group A. While numbers of L2 only utterances are small, as might be expected, proportions of L1 only and codeswitched utterances are quite similar to each other. Table 4 also gives MLU in morphemes (calculated over all the children's utterances) for each of the categories. MLU is inflated here because all single word utterances were removed from the data; however this permits comparison of single language utterances with codeswitched utterances, which by definition must contain more than one word. MLU is longest for the codeswitched data, suggesting that it is these utterances that allow the children to most fully demonstrate their linguistic ability.

Table 3

Tokens and types and distribution across languages

Group	Total tokens in code switched utterances	Total types in code switched utterances	Mirpuri tokens	Mirpuri types	English tokens	English types
A (n=6)	372 (100%)	138 (100%)	287 (77.15%)	99 (71.74%)	85 (22.85%)	39 (28.26%)
B (n=11)	843 (100%)	240 (100%)	713 (84.58%)	191 (79.58%)	130 (15.42%)	49 (20.42%)

Table 4

Mean length of utterance for Group A by utterance type (single word utterances excluded)

Analysis	All	L1 only	L2 only	CS
Utterances	87	41 (47%)	8 (9%)	38 (44%)
MLU	4.161	4.000	3.625	4.447

Table 5 gives the same figures for the older Group (B). Here there was only one L2 utterance recorded. The proportion of L1 utterances to CS is slightly higher than for the younger group. However, since there are so few L2 utterances, the older children do not appear to be losing their L1 in favor of L2, despite wider exposure to English through nursery than the younger group. As would be expected with any group of children, Group B, the older group, produced more utterances, consisting on average of more words and morphemes than the younger group, Group A.

Table 5

Mean length of utterance for Group B by utterance type (single word utterances excluded)

Analysis	All	L1 only	L2 only	CS
Utterances	209	118 (56%)	1 (1%)	90 (43%)
MLU	4.086	3.805	2.000	4.478

Table 6 shows the items that were codeswitched to English, with thematic roles assigned where appropriate, for each group. In many cases, as expected, these were content words, realizing thematic roles assigned by the verb. However English nouns also appeared in the verb phrase, accompanied by a Mirpuri operator (see further discussion of this form later).

Table 6

Intrasentential codeswitching categorized by thematic role

Thematic role	Group A (n=6)		Group B (n=11)	
	Number of children code switching this item	Total number of items of this type in sample	Number of children code switching this item	Total number of items of this type in sample
Agent	3	11	5	6
Patient*	5	26	11	43
Goal	1	1	1	1
Location	3	5	5	7
English Noun + 'Mirpuri operator (<i>kar/mar</i>)	4	10	11	25
Lexical verb	1	2	7	10

*Patient and theme are not differentiated

3.3***Examples of codeswitching***

To illustrate the patterns that emerged, there follows some examples of codeswitched utterances in the data. The frame language is identifiable as Mirpuri on the grounds of firstly, word order (SOV), and secondly late system morphemes signaling agreement between subject nouns and various verb phrase components which are all Mirpuri. There is no evidence at this stage in the children's development of English (SVO) word order, or English late system morphemes.

- (8) Codeswitching the agent and patient to English in a Mirpuri frame (26 examples from 5/6 children):

Target:

Jena gend sut -ena pija
 Man ball throw -ing+MALE is+MALE
 '(the) man is throwing (the) ball'

Child's sentence:

DADDY FOOTBALL *sat -an laga*³
 Daddy football throw -will about to
 'daddy about to throw (the) ball'

³ The meaning and grammatical status of *laga* here is unclear: grammars suggest 'about to' or 'starting to.'

(9) Codeswitching the agent, patient and goal to English in a Mirpuri frame:

Target:

Mura gudi ki to -na pija
 Boy doll to wash -ing+MALE is+MALE
 '(the) boy is washing (the doll)'

Child's sentence:

BOY BABY *ki* BATH *vitth baja*
 Boy baby to bath in put+MALE
 '(the) boy put (the) baby in (the) bath'

Comment: The child uses three English words and three Mirpuri words for this utterance. However, the Mirpuri frame conforms to Mirpuri syntactic and grammatical constraints. The object 'baby' comes before the verb, and 'in' is a postposition.

(10) Codeswitching the location to English in a Mirpuri frame:

Target:

Kuri bet -i vi
 Girl sit -ing+FEMALE is+FEMALE
 '(the) girl is sitting'

Child's sentence:

CHAIR *upr bet -i vi eh*⁴
 Chair on sit -ing+FEMALE is+FEMALE is
 '(she) is sitting on (the) chair'

Comment: The child uses "chair" instead of the Mirpuri noun *kursi*. Again, the utterance conforms to Mirpuri syntactic and grammatical constraints (Mirpuri is the frame language).

(11) Codeswitching involving a compound verb: the main element is an English noun and the operator (generic verb with contact) is Mirpuri. Sentence structure is Mirpuri:

Target:

Mura tusvir bana -na pija
 Boy picture make -ing+MALE is+MALE
 '(the) boy is drawing a picture'

Child's sentence:

Eh PEN mar -na
 Is pen DO+CONTACT -ing+MALE
 '(he) is doing (a) pen'

⁴ *eh* is actually the Punjabi auxiliary form *is*. Duplication of the auxiliary in Punjabi occurs a number of times in the data and it is not clear why. Note that *eh* is not marked for gender.

Comment: This child uses the noun + operator construction in place of the lexical verb *bana* (make) or *lik* (write). However, the child correctly selects *mar* (do + contact) retaining some of the semantic information of the lexical verb (contact of the pencil with the paper). The male gender agreement with the (omitted) subject of the sentence is maintained by the inflection of the operator.

- (12) Codeswitching involving a compound verb: the main element is an English verb, and the operator is Mirpuri:

Target:

Jena at suka -na pija
 Daddy hands dry -ing+MALE is+MALE
 'Daddy is drying his hands'

Child's sentence:

DADDY *tolija nal at WASH kar -na pija*
 Daddy towel with hands wash do -ing+MALE is+MALE
 'Daddy is doing washing (his) hands with (a) towel'

Structures of the type illustrated in (11) and (12) above are described by Romaine (1995) for Panjabi/English codeswitching, and this appears to be a common feature in contact situations involving Indian subcontinent languages. Romaine also notes that there is a semantic distinction in Panjabi between two possible operators, *hona* and *karna*, with stative verbs not occurring with *hona*. This does not map on to the Mirpuri distinction however, where contact/noncontact appears to be the main contrast involved (and where the "contact" verb *mar* means 'hit' when it occurs as a lexical verb). Romaine assigns a grammatical description to these compound verb forms in terms of government and binding theory. From the point of view of the 4M model used here, the operator and auxiliary morphemes are of the late system type, ensuring gender agreement with a subject noun phrase outside the maximal projection of the verb phrase. What is of interest is that these forms are adopted by children at a young age and this may be an example of codeswitching as a child's first language, as opposed to codeswitching that is developed by mature speakers.

3.4

Frequency of lexical items

All the data from the six Band A Mirpuri speaking children were analyzed. The children used a total of 372 word tokens which accounted for 138 different word types. Of these types, 39 were English words, many of which appeared only once in the corpus. A full list of these English words is given in Appendix 1. In order to look for evidence of an emerging preference for English or Mipuri we focused on the most common items. Cross language comparisons were made for each lexical item that occurred at least 1% of the time (in either language) across the whole corpus.

Table 7 documents items where the word was expressed at least once in the sample in English, and which contributed more than 1% of total words used.

Table 7

Frequency of Mirpuri versus English items for Group A and percentage of total vocabulary tokens

Item	Item	L1 Mirpuri	L2 English	Ratio L1:L2
<i>gend</i>	ball	0	8 (2.15%)	0:8
<i>mura</i>	boy	10 (2.69%)	7 (1.88%)	10:7
<i>aba</i>	daddy	4 (1.08%)	7 (1.88%)	4:7
no equivalent in Mirpuri	cake	-	5 (1.34%)	-
<i>charu</i>	brush	0	4 (1.08%)	0:4
no Mirpuri equivalent	phone	-	4 (1.08%)	-
<i>upr</i>	up / on	4 (1.08%)	3 (0.81%)	4:3
<i>siri</i>	ladder	4 (1.08%)	3 (0.81%)	4:3
<i>kelna</i> (play-ing + male) / <i>kelni</i> (play-ing + female)	playing	10 (2.69%)	2 (0.54%)	10:2
<i>umi</i>	mummy	6 (1.61%)	2 (0.54%)	6:2
<i>pija</i> (is + male) / <i>pi</i> (is + female) / <i>vi</i> (is + female) / <i>eh</i> (is)	is	45 (12.10%)	1 (0.27%)	45:1
<i>kuri</i>	girl	19 (5.11%)	1 (0.27%)	19:1
<i>gudi</i>	doll	4 (1.08%)	1 (0.27%)	4:1
<i>pul</i>	flower	4 (1.08%)	1 (0.27%)	4:1

Ball and *brush* here appear to be functioning as loan words, as the Mirpuri equivalents are never used by the children. *Cake* and *telephone* are also clearly loan words as there is no available Mirpuri equivalent. Of the rest, the English usage tends to be low in comparison with Mirpuri, except for *mural boy*, *aba/daddy* (where the English form is the more frequent), *upr/up* and *siri/ladder*.

Group B children used a total of 843 word tokens, and 246 word types, of which 43 were English (see Appendix 2). Table 8 documents English and Mirpuri usage for items that occurred in English on at least one occasion, and which contributed more than 1% of words used.

Here, *ball* and *brush* again seem to be functioning as loan words. The other three items occur in English only infrequently and *boy* is now more frequently expressed by the Mirpuri word, *mura*. The frequency of codeswitching as we have seen does not appear to go down over time, but the frequency with which any individual English word is used does go down. There is then very little evidence of language shift to English in these children at this stage. Only a small number of English words are used with any degree of frequency by children in both age bands. Although more data would be needed to establish this, with some analysis of patterns of integration, it would seem that these may be loan words.

Table 8

Frequency of Mirpuri versus English items for Group B and percentage of total vocabulary tokens

Item	Item	L1 Mirpuri	L2 English	Ratio L1:L2
<i>gend</i>	ball / football	0	23 (2.73%) 20 + 3	0:23
<i>charu</i>	brush	0	13 (1.54)	0:13
<i>kuri</i>	girl	33 (3.91)	5 (0.59%)	33:5
<i>pija</i> (is + male) / <i>pi</i> (is + female) / <i>vi</i> (is + female) / <i>eh</i> (is)	is	144 (17.08%)	1 (0.12%)	144:1
<i>mura</i>	boy	24 (2.85%)	1 (0.12%)	24:1

4 Discussion

The following discussion takes each of the research questions in turn:

4.1

Does Mirpuri-English codeswitching occur in the language of preschoolers with a Mirpuri-speaking home background?

As can be seen from Table 2, every child produced utterances containing intrasentential codeswitching ($n = 17$), and just over 40% of multiword utterances in the sample contain an intrasentential codeswitch. This frequency suggests that codeswitching is the norm for children in both the age groups looked at. While the number of such codeswitches is higher for the older group of children, the overall percentage remains almost the same. The proportion of use of English versus Mirpuri items remains very similar for the two age groups; however, the majority of items are still drawn from Mirpuri (Table 3), demonstrating that Mirpuri is clearly the dominant language for these children. Moreover, using criteria based on word order and morpheme type, Mirpuri is the matrix language in codeswitched utterances.

4.2

If intrasentential codeswitching does occur, is this subject to grammatical constraints, such that it conforms, even with young children, to the MLF 4M model?

There is good evidence that the instances of codeswitching observed here conform to the MLF 4M model. Table 6 shows that codeswitching results in the use of the embedded language (English) to express a range of thematic roles, most commonly patient, location and agent, as would be expected according to the model. Codeswitching the lexical verb also occurs and is more common in the older group. Word order is not disrupted or altered by the use of English items, and conforms to that of the matrix language. English items are incorporated straightforwardly, except where there is verbal involvement. Subject-verb gender agreement is retained (marked by the verb present

progressive suffix and auxiliary verb form), even where the subject is an English item. Where codeswitching occurs in the verb phrase this frequently involves a compound verb structure, as described by Romaine (1995), (see examples (11) and (12)): a generic verb carries gender agreement morphemes. Essentially, we get a structure like “do pen” for *write*, or “do wash” for *wash*, with the Mirpuri equivalent of *do* carrying inflections to mark subject-verb gender agreement. The model would predict that these “late system” morphemes should be expressed through the matrix language, and this is the case here (interestingly, the retention of the *mar/kar* distinction here also permits a semantic contrast unique to Mirpuri, that of contact/noncontact, to be expressed in the codeswitched sentence).

4.3

In terms of frequency of English items embedded into utterances otherwise spoken in Mirpuri, is there any evidence of a wholesale language shift towards English?

The frequency data demonstrated that only two English lexical items appear to have become firmly established in the language of these children (*ball* and *brush* for both groups), and a further two (*cake* and *telephone*) are expressed in English as no Mirpuri equivalent is available. The fact that *ball* and *brush* never occur in L1 and occur in both groups is evidence to support the view that the L1 equivalents are either unknown to the children or very low frequency words in their vocabulary.

In other cases proportions of L1 to L2 lexical item use for the same referent are distributed more evenly. This suggests that the L1 and L2 lexical items are being treated as synonyms. Examples of these are *mura*, ‘boy’ (10:7 L1:L2 usage), *aba*, ‘daddy’ (4:7), and *siri*, ‘ladder’ (4:3) for Group A. Group B is more polarized, with only *kuri*, ‘girl’ (33:5) as an example in the top 1% of items. For a further item, *upr*, ‘up’ or ‘on’, used to a similar extent in both languages by Group A, it is possible that there may be some confusion and merging of the two forms as a result of their phonetic and semantic similarity.

It should be noted that frequency of word use will have been affected by the assessment picture stimuli which limit the children’s possible responses. Comparison with a wider corpus of spontaneous utterances would be useful here.

4.4

What are the implications for norms of bilingual language acquisition, and for language assessment?

The children from this population are different from their monolingual Pakistani peers and their English monolingual peers. For the children in this study, codeswitching appears to be the norm for this stage of language acquisition. The MLUs for code-switched utterances in both groups are higher than L1 utterances, L2 utterances or total utterances in all cases (Tables 4 and 5). This suggests that the codeswitched utterances represent the more complex utterances used by bilingual children at this stage, and to disregard them would give a misleading picture of any individual child’s language development. Furthermore this is definitely not a case of random mixing of items within a sentence. The codeswitched utterances showed a high degree of conformity to the MLF 4M model and a degree of grammatical sophistication commensurate with monolingual

children at a similar age. In fact it is possible that a lack of codeswitching in children in this population may be an indicator of language delay or intrinsic disorder.

There were no major differences between the two age groups. Older children responded to more items and with longer utterances, as would be expected. Overall percentage of codeswitched utterances between the two groups was very similar (Group A, 44%; Group B, 43%). The older children had more switches within the verb phrase, involving a compound verb structure with the operator in Mirpuri. Frequency data however indicated increasing polarization of the two languages in terms of choice of lexical items. It is possible that children become more aware of two independent language systems at this age. It will be important to look at later stages of development. We would expect use of English to become more distinct and highly developed, with a possible later shift to English language dominance. Children may also be expected to recognize nuances in the range of codeswitching in the adults around them, and may adopt a more “pure” Mirpuri in response to some circumstances, including linguistic domain. An alternative scenario is that the community as a whole may be adopting a new variety of the language (influenced by L2) as proposed by Martin in relation to Panjabi/English speaking children, (Martin, et al. 2003, p.261), in which case the children’s L1 would move closer to this variety over time.

5 Conclusion

This study looks at a previously undocumented group of bilingual speakers, that of the Pakistani heritage population in Rochdale, U.K. Furthermore, where the majority of studies of codeswitching have involved data collected from adults, the data here is from young preschool children. The aim has been to look at emerging codeswitching, rather than viewing use of more than one code as a temporary “mixing” stage. The responses of the 17 children, all with L1 Mirpuri and beginning to acquire L2 English, clearly demonstrated the use of codeswitching. Furthermore, this codeswitching was subject to grammatical constraints and conformed to the MLF/4M models. Frequency data, together with the more complex verb phrase examples of codeswitching, suggests that we have an emerging codeswitched variety here and not a pattern of shift towards monolingual English. It is also clear that, at this stage in their language acquisition, use of codeswitched utterances is an important marker of the child’s linguistic ability. This has major implications in terms of establishing norms for language acquisition in this group of children, and for the types of language assessment that might be carried out by speech and language therapists.

The study has been hampered by a lack of linguistic description of the language or dialect used by the children (Mirpuri) and information on how it relates to the other varieties used by this population. Furthermore, no systematic sociolinguistic account exists for this community, so the patterns of usage of the three languages are largely unknown. It is known that this community experiences significant social deprivation, such that many of its members live in “Sure Start” areas. Given these facts, there is an urgent need for studies addressing the Mirpuri language and the sociolinguistic make-up of the area, plus further work on language acquisition. Such work will inform

attempts to improve educational outcomes and address language difficulties where they might arise.

manuscript received: May, 2005

revision received: December, 2005

manuscript accepted: December, 2005

References

- ABUDARHAM, S. (1987). Fact and fiction. In S. Abudarham (Ed.), *Bilingualism and the bilingual: An interdisciplinary approach to pedagogical and remedial issues* (pp. 15–34). Windsor: NFER-NELSON.
- AUER, P. (1998). Introduction: Bilingual conversation revisited. In P. Auer (Ed.), *Codeswitching in conversation: Language interaction and identity* (pp. 1–24). London: Routledge.
- AZUMA, S. (1993). The frame-content hypothesis in speech production: Evidence from intra-sentential codeswitching. *Linguistics*, **31**, 1071–1093.
- BENTAHILA, A., & DAVIES, E. E. (1995). Patterns of codeswitching and patterns of language contact. *Lingua*, **96**, 75–93.
- BHARDWAJ, M. R. (1995). *Colloquial Panjabi: A complete language course*. London: Routledge.
- BHATT, R. M. (1997). Codeswitching, constraints, and optimal grammars. *Lingua*, **102**, 223–251.
- BOUMANS, L. (2001). Searching for syntactic explanations of codeswitching. *Linguistics*, **39**, 437–453.
- BOZTEPE, E. (2003) Issues in codeswitching: Competing theories and models. *Working Papers in TESOL and Applied Linguistics Web Journal*, **3**, 1–27. [<http://www.tc.columbia.edu/tesolalwebjournal/Boztepe.pdf>].
- BRICE, A., & ANDERSON, R. (1999). Code mixing in a young bilingual child. *Communication Disorders Quarterly*, **21**, 17–22.
- BRITISH BROADCASTING CORPORATION. (2004a) *BBC Asian Network: Mirpuri*. Accessed on 26-NOV-2004. [http://www.bbc.co.uk/asiannetwork/presenters/showpage_mirpuri.shtml]
- BRITISH BROADCASTING CORPORATION. (2004b). *BBC Asian Network: Zarina Khan biography*. Accessed on 08-APR-2004. [http://www.bbc.co.uk/asiannetwork/presenters/biog_zarina_khan.shtml]
- CHENG, K. Y. (2003). Codeswitching for a purpose: Focus on preschool Malaysian children. *Multilingua*, **22**, 59–77.
- CHIOCCA, E. M. (1998). Language development in bilingual children [Jan/Feb]. *Pediatric Nursing*, **24**, 43–47.
- CUNNINGHAM-ANDERSSON, U., & ANDERSSON, S. (2004). *Growing up with two languages: A practical guide*. London: Routledge.
- DRIVING STANDARDS AGENCY, (2005) *Special needs and the theory test*. Accessed on 01-MAR-2005. [<http://www.dsa.gov.uk/Cars.asp?cat=92>]
- GARDNER-CHLOROS, P., & EDWARDS, M. (2004). Assumptions behind grammatical approaches to codeswitching: When the blueprint is a red herring. *Transactions of the Philological Society*, **102**, 103–129.
- GLASS, N. (2001). What works for children—the political issues. *Children and Society*, **15**, 14–20.
- GOODZ, N. S. (1989). Parental language mixing in bilingual families. *Infant Mental Health Journal*, **10**, 25–44.
- HEREDIA, R. R., & ALTARRIBA, J. (2001). Bilingual language mixing: Why do bilinguals codeswitch? *Current Directions in Psychological Science*, **10**, 164–168.

- INTERNATIONAL PHONETIC ASSOCIATION, (1996). *International Phonetic Alphabet*. Accessed on 13-FEB-2005. [<http://www2.arts.gla.ac.uk/IPA/ipachart.html>]
- ISAAC, K. (2002). *Speech pathology in cultural and linguistic diversity*. London: Whurr.
- JISA, H. (2000). Language mixing in the weak language: Evidence from two children. *Journal of Pragmatics*, **32**, 1363–1386.
- KIRKLEES METROPOLITAN COUNCIL EDUCATION SERVICE, & DEWSBURY COLLEGE, (2002) *Joint report visit to Azad Kashmir, Pakistan*. Accessed on 08-FEB-2005. [<http://www.kirklees-ednet.org.uk/subjects/ethnicmin/docs/azadkashmir/>]
- LANZA, E. (1997). *Language mixing in infant bilingualism*. Oxford: Oxford University Press.
- LEONARD, L. B. (1998). *Children with specific language impairment*. London: MIT Press.
- LI WEI, (2000), Dimensions of bilingualism. In Li Wei (ed.), *The bilingualism reader* (pp.3–25) London: Routledge.
- LINGUISTIC MINORITIES PROJECT. (1985). *The other languages of England*. London: Routledge and Kegan Paul.
- LIPPS GROUP (Language Interaction in Plurilingual & Plurilectal Speakers) (2000). The LIDES Coding Manual: A document for preparing and analyzing language interaction data, Version 1.1 [July, 1999]. *The International Journal of Bilingualism*, **4**, 131–270.
- MacW.HINNEY, B. (2003a). *CLAN (Computerized Language Analysis) manual*. Accessed, on 14-FEB-2005. [<http://childes.psy.cmu.edu/manuals/CLAN.pdf>]
- MacW.HINNEY, B. (2003b). *CHAT transcription manual*. Accessed on 01-MAR-2005. [<http://childes.psy.cmu.edu/manuals/CHAT.pdf>]
- MARTIN, D. R. K., KRISHNAMURTHY, R., BHARDWAJ, M., & CHARLES, R. (2003). Language change in young Panjabi/English children: Implications for bilingual language assessment. *Child Language Teaching and Therapy*, **19**, 245–265.
- MARTIN, D., KRISHNAMURTHY, R., BHARDWAJ, M., & CHARLES, R. (2004). Clarification note. *Child Language Teaching and Therapy*, **20**(3).
- MASIDLOVER, M., & KNOWLES, W. (1979). *Derbyshire Language Scheme: Rapid screening test*. Derby: Derbyshire County Council.
- MYERS-SCOTTON, C. (2002). *Contact linguistics: Bilingual encounters and grammatical outcomes*. Oxford: Oxford University Press.
- NATIONAL STATISTICS (2003). *Statistics about Rochdale*. Accessed on 29-NOV-2004. [<http://neighbourhood.statistics.gov.uk/AreaProfileFrames.asp?TID=13&AREA=Rochdale&AID=175531>]
- NATIONAL STATISTICS (2004a). *Ethnicity and identity: Population size*. Accessed, on 29-NOV-2004. [http://www.statistics.gov.uk/cci/nugget_print.asp?ID=455]
- NATIONAL STATISTICS (2004b) *Ethnicity and identity: Geographic distribution*. Accessed on 29-NOV-2004. [http://www.statistics.gov.uk/cci/nugget_print.asp?ID=457]
- NERWAL, H. (2002). *Nottingham Health Action Zone final evaluation report for touchscreen project*. Accessed on 08-APR-2004. [http://www.haznet.org.uk/hazs/evidence/local_notts-touchscreen.pdf]
- NICOLADES, E., & GENESEE, F. (1997). Language development in preschool bilingual children. *Journal of Speech Language Pathology and Audiology*, **21**, 258–270.
- PARADIS, J., NICOLADIS, E., & GENESEE, F. (2000). Early emergence of structural constraints on code mixing. *Bilingualism: Language and Cognition*, **3**, 245–261.
- PERT, S. (2000). *Derbyshire language scheme—rapid screening test: Adaptations for Mirpuri, Punjabi and Urdu* (Rochdale: Unpublished assessment adaptations).
- PERT, S., & LETTS, C. (2003). Developing an expressive language assessment for children in Rochdale with a Pakistani heritage background. *Child Language Teaching and Therapy*, **19**, 267–289.

-
- PERT, S., & STOW, C. (2001). Language remediation in mother tongue: A pediatric multilingual picture resource. *International Journal of Language and Communication Disorders*, **36**, (Supplement): 303–309.
- PFAFF, C. W. (1979). Constraints on language mixing: Intrasentential codeswitching and borrowing in Spanish / English. *Language*, **55**, 291–318.
- POPLACK, S. (1980). Sometimes I'll start a sentence in Spanish *y termino en Español*: Towards a typology of codeswitching. *Linguistics*, **18**, 581–618.
- REDLINGER, W., & PARK, T.-Z. (1980). Language mixing in young bilingual children. *Journal of Child Language*, **7**, 337–352.
- REYES, B. A. (1995). Considerations in the assessment and treatment of neurogenic communication disorders in bilingual adults. In H. Kayser (Ed.), *Bilingual speech-language pathology: An Hispanic focus* (pp. 153–182). London: Singular.
- REYNOLDS, M. (2002). Punjabi/Urdu in Sheffield: Language maintenance and loss and development of a mixed code. In P. Gubbins & M. Holt (Eds.), *Beyond boundaries: Language and identity in contemporary Europe* (pp. 145–162). Clevedon: Multilingual Matters.
- ROMAINE, S. (1986). The syntax and semantics of the codemixed compound verb in Panjabi/English bilingual discourse. In D. Tannen & J. E. Alatis, (Eds.), *Languages and linguistics: The interdependence of theory, data and application* (pp. 35–50). Washington, DC: Georgetown University Press.
- ROMAINE, S. (1995). *Bilingualism*. Oxford: Blackwell Publishing.
- SAEED, J. I. (2003). *Semantics*. Oxford: Blackwell Publishing.
- SANKOFF, D. (1998) A formal production-based explanation of the facts of codeswitching. *Bilingualism, Language and Cognition*, **1**, 39–50.
- SCOTT, A., PEARCE, D., & GOLDBLATT, P. (2001). The sizes and characteristics of the minority ethnic populations of Great Britain—latest estimates. *Population Trends*, **105**, 6–15.
- SCHMIDT, R. L. (2004). *Urdu: An essential grammar*. London: Routledge.
- SHIN, S. J. (2002). Differentiating language contact phenomena: Evidence from Korean-English bilingual children. *Applied Psycholinguistics*, **23**, 337–360.
- SPEKTOR, L., & TUTHILL, B. (2000). CLAN: Computerized language analysis: CMU.
-

Appendix 1

English words used by Group A (number of tokens in brackets)

Ball	(8)	Mummy	(2)	Hello	(1)
Boy	(7)	Pen	(2)	Is	(1)
Daddy	(7)	Play	(2)	It	(1)
Cake	(5)	Wash	(2)	Laugh	(1)
Brush	(4)	Bath	(1)	Newspaper	(1)
Chair	(4)	Blue	(1)	Paint	(1)
Cut	(4)	Book	(1)	Pencil	(1)
Phone	(4)	Climb	(1)	She	(1)
Baby	(3)	Cook	(1)	The	(1)
Ladder	(3)	Cycle	(1)	With	(1)
Stand	(3)	Doll	(1)	Write	(1)
Up	(3)	Flower	(1)		
Football	(2)	Game	(1)		
Machine	(2)	Girl	(1)		

Appendix 2

English words used by Group B (number of tokens in brackets)

Ball	(20)	Kick	(3)	Friend	(1)
Brush	(13)	Ladder	(3)	Game	(1)
Phone/ telephone	(10)	Boot	(2)	Hand	(1)
Cake	(9)	Door	(2)	Hello	(1)
Baby	(8)	Is	(2)	News	(1)
Chair	(6)	Mummy	(2)	Paper	(1)
Girl	(5)	Newspaper	(2)	School	(1)
Book	(4)	Pencil	(2)	Scooter	(1)
Machine	(4)	Shoe	(2)	Stand	(1)
Color	(3)	Bath	(1)	Throw	(1)
Cut	(3)	Boy	(1)	Toffee	(1)
Flower	(3)	Bubbly	(1)	Towel	(1)
Football	(3)	Burger	(1)	Write	(1)
Hoover	(3)	Chuck	(1)	Write	(1)
		Flower	(1)		