

© 2002 Cambridge University Press. Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing [scholarworks-group@umbc.edu](mailto:scholarworks-group@umbc.edu) and telling us what having access to this work means to you and why it's important to you. Thank you.

Citation:

Shin, S. J. (2002). Differentiating language contact phenomena:  
Evidence from Korean-English bilingualism. *Applied  
Psycholinguistics*, 23(3), 337-360.

**Differentiating language contact phenomena:  
Evidence from Korean-English bilingualism  
(Differentiating language contact phenomena)**

**Sarah J. Shin**

**Department of Education**

**University of Maryland Baltimore County**

**1000 Hilltop Circle**

**Baltimore, MD 21250, USA**

**Phone: 410-455-2485 (Office)**

**410-203-2480 (Home)**

**Fax: 410-455-3986 (Office)**

**E-mail: [shin@umbc.edu](mailto:shin@umbc.edu)**

# **Differentiating language contact phenomena: Evidence from Korean-English bilingualism**

## **Abstract**

This paper attempts to provide a reliable description of the characteristics of intrasentential language mixing produced by a group of Korean-English bilingual children, with a special focus on the distinction between codeswitching and borrowing. Making use of the inherent variability in case marking in Korean, this study employs a quantitative variationist method to determine the status of single nouns of English-origin in an otherwise Korean discourse, which constitute the largest portion of the bilingual data. Analysis of the overall bilingual data suggests that intrasentential language mixing is determined by the bilingual abilities and preferences of the speaker as well as those of the addressee. Results of the variable analysis of case marking shows that most of the English-origin objects and subjects of Korean verbs are best treated as borrowings and not as codeswitches.

## **Introduction**

There is general agreement in the theoretical studies of bilingualism that there are certain syntactic constraints on intrasentential codeswitching between any pair of languages. A significant number of studies have revealed that intrasentential codeswitching is not located randomly in the sentence, but rather that it occurs at specific points. Where much less agreement is reached is with respect to general properties of the process. Various ‘constraints’ and ‘models’ regulating intrasentential codeswitching have been proposed and tested, with the result that some cases appear to be explained by one constraint, and others by another (e.g. Poplack’s (1980) Free Morpheme Constraint, Di Sciullo, Muysken & Singh’s (1986) Government Constraint, Myers-Scotton’s (1993a) Matrix Language Frame model and Myers-Scotton & Jake’s (2000) 4-M Model, and Belazi, Rubin, & Toribio’s (1994) Functional Head Constraint). One of the major issues in this ongoing debate is the task of distinguishing between codeswitching and borrowing. Whereas borrowing is the adaptation of lexical material to the patterns of the recipient language, codeswitching is the juxtaposition of sentence fragments formed according to the internal syntactic rules of two distinct systems (Poplack, 1993).

The root of the problem lies in the status of single lexical items from one language inserted in discourse in another, a phenomenon that constitutes the greatest portion of any bilingual corpus systematically studied (e.g. Berk-Seligson, 1986; Poplack & Meechan, 1998). Single word switches pose significant problems because they are often not easily categorized as either codeswitches or borrowings. Researchers who make a distinction between borrowing and codeswitching have generally made reference to the degree of integration of the borrowed items in the base language (Romaine, 1995, p.

143). Grosjean (1982, p. 308) defines codeswitching as an other language item of any length (a word, a phrase, a sentence) which is a complete shift from the other language, whereas he defines borrowing as a word or short expression that is adapted phonologically and morphologically to the language being spoken. Grosjean admits, however, that it is not always possible to make a clear-cut distinction between borrowings and codeswitching using these criteria. Pfaff (1979, p. 297) claims that in order to fully determine the status of a word, one would have to know whether a base language equivalent existed; and if so, whether it is in use in the community, whether the individual knows it and whether he or she regards it as belonging to one language or another.

Perhaps the most systematic approach to distinguishing codeswitching from borrowing is found in the research by Shana Poplack and her colleagues. Using a large bilingual speech corpus, Poplack, Sankoff and Miller (1988) quantitatively examined the use of single English words in French contexts in five francophone regions in Ottawa. They compared attestation histories of English forms in Canadian and European French with usage frequencies in their data and various measures of integration to distinguish two types of borrowing -- nonce and established. Established borrowings typically show full syntactic, morphological, and usually phonological integration into the patterns of the recipient language and are widely used and recognized by the recipient-language monolinguals. On the other hand, nonce borrowings are identical to established borrowings in terms of linguistic integration but are not widely diffused and so are not generally used and recognized by recipient-language monolinguals. While this distinction is useful, Poplack *et al.* (1988) concede that in French-English bilingualism, it is difficult

to distinguish codeswitching from borrowing at the single-word level since there are often no morphological or syntactic criteria for determining whether a single lexical item is following English or French rules. For instance, noun morphology is most frequently null in both languages, and word order is very similar.

Later research suggests that this problem can be resolved by quantitatively examining typologically different language pairs (e.g. Ukrainian/English – Budzhak-Jones, 1998; Igbo/English – Eze, 1998; Wolof/French and Fongbe/French -- Poplack & Meechan, 1995). As Sankoff, Poplack, & Vanniarajan (1990) point out, quantitative analysis has an important advantage -- while it is often impossible to determine the status of individual lexical items in isolated sentences, one can arrive at a reasonably reliable solution based on the analysis of *all* of the tokens in a given category. For example, focusing on points within the sentence where the morphosyntactic structures of Ukrainian (a fusional language) and English (an analytic language) conflict, Budzhak-Jones (1998) compares the distribution of all of the English-origin nouns with overt Ukrainian morphology in otherwise Ukrainian discourse with that of native Ukrainian nouns to show that the English-origin nouns do not behave as if they were English. Instead, they replicate the patterns of behavior of monolingual Ukrainian nouns with respect to gender assignment, modifier-noun agreement, and inflectional variability and are thus borrowings rather than codeswitches. Similarly, Poplack and Meechan (1995) show that single French-origin nouns in otherwise Wolof and Fongbe discourse pattern like monolingual Wolof and Fongbe nouns with regard to their modification structures and are thus borrowings.

While the existing research that draws on the variationist paradigm has contributed significantly to our understanding of the structural properties of languages in contact, most of this work has involved adult bilingual speakers and it is unclear whether and how the quantitative methods outlined by Poplack and her colleagues can be applied to children's bilingual speech. The current study is an attempt to contribute to filling this gap in the literature. Another motivation for this study is that there are currently relatively few studies that have investigated the development of formal and functional principles that govern bilingual children's codeswitching. Most investigations of the structural properties of bilingual children's language mixing have examined the frequencies of different syntactic categories occurring as single-item insertions in mixed utterances (e.g. Deuchar & Quay, 1998; Köppe, 1996; Köppe & Meisel, 1995; Lanza, 1992, 1997b; Meisel, 1994; Paradis, Nicoladis, & Genesse, 2000; Vihman, 1985, 1998). For example, grammatical morphemes have been found to be usually more restricted in where they can be inserted in a sentence than content morphemes such as nouns (Köppe & Meisel, 1995; Redlinger & Park, 1980).

Some of these studies have also dealt with the question of whether or not bilingual children's code-mixing obeys the same grammatical constraints proposed for adult bilingualism. For example, Meisel (1994) examined the mixed utterances of two French-German bilingual children in reference to the Government Constraint proposed for adult bilinguals and concluded that the children's mixed utterances adhere to a modified version of this constraint. Similarly, Lanza (1997a), Vihman (1998), and Paradis et al. (2000) have examined bilingual children's code-mixed utterances in reference to the

Matrix Language Frame (MLF) model and concluded that the structure of the children's language mixing generally followed modified versions of this adult model.

However, a serious limitation of these and most other childhood bilingualism studies is that they simply assume all materials of other language origin to be codeswitches, giving no regard to the possibility that some of them may actually be borrowings. Meisel (1994), in his argument that nouns and noun phrases are the most switched elements in children's bilingual discourse, attempted to explain apparent exceptions to this pattern (namely, that verbs are the most switched items in Panjabi/English bilingualism (Romaine, 1995, p.131)) by stating that these exceptions may in fact be borrowings rather than codeswitches. This ad-hoc categorization of certain lexical items as borrowings and others as codeswitches is problematic because it does little to improve our understanding of these different language contact phenomena. The current study attempts to address this problem directly by examining children's patterns of codeswitching in light of the quantitative principles Poplack et al. have articulated for adult codeswitching. More specifically, the current study investigates children's language mixing involving Korean and English, a typologically different language pair.

In addition, bilingual children's speech productions have often been studied without reference to the sociolinguistic contexts in which they occurred, hence making it impossible to analyze the possible factors underlying any codeswitching behavior (De Houwer, 1995). For instance, can children's intransentential language mixing be seen as a function of language proficiency and/or preference of the speaker and of the interlocutor? Or, can it be seen as a function of the setting such as the number of same language speakers present? The current study attempts to provide a reliable description of the

structure of the mixed language of a group of Korean-English bilingual children while taking these and other sociolinguistic factors into account. More specifically, it investigates peer discourse among Korean children in a mostly monolingual English-speaking classroom.

### **The present study**

In the current children's bilingual corpus, many of the English-origin nouns in otherwise Korean contexts are ambiguous as to their degree of integration into the linguistic structures of Korean. This is because many of them appear bare without any indication of linguistic integration such as noun modification or case markers, as shown in (1).

(1) story **mantul-eyatoy**.<sup>1</sup>

make-must

([We] need to make a story.)

Some scholars, including those that have specifically investigated adult Korean-English codeswitching, have simply assumed such nouns to be codeswitches, and have reported them as violations of the grammars of the languages in contact (e.g. Bentahila & Davies, 1991; Bokamba, 1988; Choi, 1991; Myers-Scotton, 1993a; Yoon, 1992). In particular, examples like this are cited as violations of the equivalence constraint (Sankoff & Poplack, 1981), which states that codeswitching should only occur at boundaries common to both languages and should not occur between any two sentence elements that are ordered differently. Since a typical word order in Korean is SOV while that in English is SVO, any English-origin words in object positions of Korean verbs are

in apparent violation of the grammars of these languages. The fact that many single words of one language origin lack indication of linguistic integration into the recipient language has led some researchers to conclude that codeswitching and borrowing are either operationally indistinguishable or undifferentiated by the bilingual speaker (Bentahila & Davies, 1991; Myers-Scotton, 1993a). In this study, however, I show that a quantitative variationist method (e.g. Adalar & Tagliamonte, 1998; Budzhak-Jones & Poplack, 1997) can effectively determine the status of these ambiguous items. This method is applied to the inherent variability in case marking in Korean, which I describe in more detail below.

### **Postpositional case markers in Korean**

Korean is a typologically SOV language with relative freedom of word order (Kim, Y., 1997; Martin, 1992; O'Grady, 1991). Such freedom of word order is possible in Korean because the grammatical roles of each constituent in a sentence are marked by postpositional case markers/particles, as shown in (2). Subjects in Korean take the nominative marker, *-ka* (following a vowel) or *-i* (following a consonant), while direct objects in Korean are followed by the accusative marker, *-(l)ul*. Indirect objects in Korean take a dative marker, *-hanthey*, or *-eykey*, while the topic marker, *-(n)un*, topicalizes nominative and accusative nouns by replacing their respective case markers, as shown in (3). In addition, a class of particles called 'delimiters' conveys information carried in other languages by articles, adverbs, prosodic elements, or word order. Delimiters carry quantificational information and do not make the structural distinctions typically associated with case (O'Grady, 1991). In fact, as (4) and (5) show, the same

delimiter used with a dative particle can appear with a subject in lieu of the nominative marker, *-i*.

Korean is classified as a null subject language. However, not only the subject of a sentence but also any or all of the nominal arguments of a predicate (i.e. verbs, adjectives, and the copula), and even the predicate itself may be dropped in Korean. In addition, case marking in Korean is optional and case markers attached to NP's may be dropped. Case particles can be dropped with different frequencies: accusative markers are dropped quite frequently while nominative markers are dropped less frequently. The inherent variability in case marking in Korean is crucial in subsequent analysis of English-origin nouns embedded in Korean discourse.

(2) **John-i Sue-hanthey kong-ul cwu-ess-ta.**

John-NOM Sue-DAT ball-ACC give-PAST-DECL

‘John gave a ball to Sue.’

(3) **John-un Sue-hanthey kong-ul cwu-n-ta.**

John-TOP Sue-DAT ball-ACC give-PRES-DECL.

‘As for John, he gives a ball to Sue.’

(4) **John-i Sue-hanthey-man kong-ul cwu-n-ta.**

John-NOM Sue-DAT-DEL (‘only’) ball-ACC give-PRES-DECL

‘John gives the ball only to Sue.’

(5) **John-man Sue-hanthey kong-ul cwu-n-ta.**

John-DEL (‘only’) Sue-DAT ball-ACC give-PRES-DECL

‘Only John gives the ball to Sue.’

## **Method**

### ***Subjects***

The data for this study were collected from tape-recorded natural conversations of 12 (six male and six female) Korean-English bilingual children. These subjects were selected on the basis that they were all in one first grade class<sup>2</sup> of 27 students and had Korean as their native language. Each child's name, sex, and age at the beginning of the fieldwork period (May, 1995) and whether or not he/she is enrolled in pullout ESL and bilingual Korean/English classes are listed in Table 1. While the subjects differed in terms of place of birth, age of immigration, and proficiencies in Korean and English, all twelve children learned Korean in the home and were systematically exposed to English when they entered school. In this sense, they are 'successive bilinguals' and are distinguished from children who acquire two languages more or less simultaneously (i.e. 'simultaneous bilinguals'). Besides Matthew, Joshua, Abel, and Kyung who were born in the U.S. and Gina who was born in Argentina, all other children were born in Korea and have subsequently moved to the U.S. with their families. Kwon's family had moved to Mexico soon after his birth and came to the U.S. when he was four years old. Except for David whose first contact with English was in this first grade class due to his recent arrival in the U.S., all eleven children attended English-speaking kindergartens (and in some cases, also preschools) in the U.S. before becoming first graders.

*(Insert Table 1 about here.)*

### ***The Teacher***

Mrs. Kim, the homeroom teacher, immigrated to the United States at the age of seven with her family from Korea. Since then, she has received her elementary, secondary, and college education in the States. While her ability in Korean has not progressed much since moving to America, Mrs. Kim can nevertheless carry on a simple conversation in Korean and speaks Korean with the Korean parents of her students. There is no trace of Korean accent in her English, but some of her Korean students attempted to speak to her in Korean in the beginning of the school year. Mrs. Kim reported having specifically instructed her Korean students not to speak to her in Korean out of consideration for the non-Korean students in her class. The fact that ten of her twelve Korean students had the opportunity to speak Korean in the daily pull-out bilingual Korean/English class also led her to insist on English as the main language in her classroom. However, although she did not allow her Korean students to address her in Korean, Mrs. Kim did not attempt to prevent them from speaking Korean among themselves.

### ***Data Collection***

Participant observation procedures were employed by the author, a bilingual Korean/English speaker, who took on the role of a teacher's assistant in the classroom. A corpus of spontaneous speech of the twelve Korean children was collected by audio-recording during a two-month fieldwork period. As shown in Table 2, based upon the homeroom teacher's evaluations on students' language proficiency, the twelve Korean

students were paired into six groups, with the two students in each group having comparable proficiency in English and Korean.

*(Insert Table 2 about here.)*

Audio-recordings were made in three situations:

- 1) *storytelling*: telling a spontaneously created story, opinion, or views to the partner based upon an activity in class.
- 2) *math*: involved counting numbers of some sort, such as in buying and selling toy goods in a "store" activity, sorting and counting different plastic shapes, measuring how far a snail travels in a given amount of time.
- 3) *play*: as part of the "Learning Center" in which children are free to play educational games with one another (e.g., various board games, wooden blocks, and jigsaw puzzles).

A light-weight wireless radio microphone, worn by one of the two children being recorded in a pair, was used to record speech from any part of the classroom while allowing children to freely move around as usual. The recordings for each pair for each activity type lasted between 20 and 75 minutes, yielding a total of approximately 10 hours of recorded speech.

### **The Data**

In general, language mixing within the clause or sentence boundary is rare when considered as a proportion of the entire corpus – a little over 3% (or 282 instances) of the

total 8447 utterances in the spontaneous speech corpus contained any kind of intrasentential language mixing. A similar pattern was reported by Poplack et al. (1988), which, despite a large database for a study of potential English loanwords into French, found only less than 1% of the total verbal output to be of English-origin. Low rates of intrasentential language mixing are also documented in studies that specifically examined children's use of various language pairs (e.g. Köppe & Meisel, 1995; Moffat & Milroy, 1992).

A closer examination of the data reveals that language mixing is unevenly distributed across subjects -- approximately 80% of the total intrasentential language mixing was produced by four of the twelve subjects (i.e. Yooni and Grace [pair 3] and David and So Hee [pair 6]). Among the twelve Korean-English bilingual subjects, these four are particularly proficient in Korean. David and So Hee, in particular, produced a large number of English numeral + noun constructions embedded in Korean discourse during their math activity, giving rise to their relatively high mixing rates. Although Kwon and Gina are also highly proficient speakers of Korean, their being paired with less proficient Korean speakers (i.e. Joshua and Kathy) may have influenced the low rates of language mixing found in their conversations. The six subjects in pairs 1, 2 and 4 spoke almost exclusively in English and had little or no language mixing.

Figure 1 shows the percentage of intrasentential language mixing each subject produced in each of the three activities. 100% on the graph represents the total number of all utterances each child produced in a given activity. A two-way analysis of variance (ANOVA) with subject pairings and activity types as two factors revealed that pairings was a significant effect on the amount of intrasentential language mixing at 0.05 level (*p*-

value= $5.73 \times 10^{-6}$ ), whereas activity types was not ( $p$ -value=0.1373). Recall that subject pairings were decided on language proficiency grounds (i.e. two students in each pair having comparable proficiency in both Korean and English). This result implies that language proficiency may have been partially responsible for the amount of language mixing produced by each child. Had the twelve subjects been paired differently (e.g. a child who speaks better English paired with a child who speaks better Korean), the amount of language mixing produced by each child could have been different.

*(Insert Figure 1 about here.)*

### **Language proficiency and preference**

When we examine the children's conversations more carefully, we indeed find that proficiency in Korean influences the amount of language mixing produced (see further, Shin & Milroy, 1999; 2000). One can see a clear example of this in the conversation between So Hee and David, the two most proficient Korean speakers among the twelve children. In excerpt (1) below, David and So Hee are paired to do a "store" activity, in which they are required to each draw three items inside a picture of an empty store and sell the items to the partner. Notice that a large majority of the conversational sequences are in either Korean or mixed Korean and English. When we examine the English words used in otherwise Korean contexts, they are mostly either numerals (plus nouns) as in line 6 or single nouns as in line 12 (more detailed descriptions of the mixed categories follow in subsequent sections). The large number of English-origin numeral + noun constructions found in otherwise Korean contexts in this sequence may have



TWENTY CENTS

(Twenty cents. As for three of them, it's twenty cents.)

7 (3.0) AH: TEN CENTS/

8 David: *nay* TURN/

my

(My turn.)

9 (3.0) I WANT TO DRAW (1.0) CUP/

10 UM I MEAN (1.2) CUP/

11 (2.0) WHAT (0.5) WHAT I-/

12 CUP *kulil-kkeya* *na* *nun/*

draw will I TOP

(I'm going to draw a cup.)

13 (11.0) (softly) *daman/*

(nonsense word)

14 (9.0) (loudly) CRAYON/

15 (David draws a cup in the "store") (16.7) *ike* *nun/*

this TOP

(As for this...)

16 (2.0) TWENTY FIVE CENTS/

17 (4.0) FIFTY CENTS/

18 (5.5) (softly) FIFTY CENTS/

19 (2.0) (loudly) HOTDOG/

20 (8.0) (softly) IT'S/  
 21 (2.0) (softly) UM IT'S/  
 22 (9.0) (abruptly and loudly) TEN CENTS/  
 23 So Hee: *ya an-toy/*  
           hey not work  
           (Hey, no!)  
 24 *ne say kay kulyesse?!*  
       you three CLASS drew  
       (Did you draw three things?)  
 25 *Say kay kulyesse?!*  
       three CLASS drew  
       (Did you draw three things?)  
 26 David: ONE TWO THREE  
 27 So Hee: OKAY *nay chalay ya/*  
                   my turn COP  
           (Okay. It's my turn.)  
 28 David: *ta kulyesse/*  
           all drew  
           (I'm finished drawing.)  
 29 So Hee: (3.0) GOOD THING *ya (unintelligible)/*  
                                   hey  
                                   (Good thing. Hey. (unintelligible).)  
 30 *ca ne mence BUY hay naykkel*

now you first do mine

(Now. You buy mine first.)

31 [WHAT DO YOU WANT/

32 David: [UM/

33 UH CRAYON?/

34 So Hee: (unintelligible) AND THIS IS PIZZA./

35 THIS IS UM WHOLE PACK IS FIFTY CENTS/

36 NO [FIFTEEN CENTS/

37 David: [ICE CREAM/

38 [ICE CREAM/

39 So Hee: [ICE CREAM?/

40 THIS IS FIFTY CENTS/

41 David: FIFTY CENTS?/

42 (4.0) (SOFT VOICE) FIVE TEN (0.7) FIFTEEN TWENTY-/

43 So Hee: (2.0) *ike lo hay to toy/*

*this with do also work*

(You can do it with this also.)

44 ONE TWO=/  
45 David: (loud voice) =a na an hay an hay

ah I not do not do

*halkkeya na kunyang/*

do will I simply

(Ah. I'm not going to do it! I'm not going to do it! I'll just do it!)

46 (1.7) TEN (1.0)

47 So Hee: I MEAN FIVE TEN TWENTY=/

While it seems quite clear that language proficiency plays a role in determining children's intrasentential language mixing behavior, language abilities alone do not explain why a proficient Korean speaker like Jae code-mixes very little. An examination of Jae's conversational data suggests that language preference also affects child language mixing patterns. Consider excerpt (2) below where Jae and Abel are given a snail to study. Their assignment is to measure the length of the snail, identify the different body parts, and measure how far it travels. In line 10, Abel initiates an extrasentential switch into Korean saying that one can eat the snail shell because it is hard. It is possible that he may have meant that one *cannot* eat the shell because it is hard, but in any event, Jae corrects Abel by saying that one cannot eat it. Abel then responds with an additional piece of information, namely that the hardness of the shell could cause one to die in case of ingestion. In line 15, the Korean sequence ends when Jae explicitly directs Abel to switch to English. This directive effectively ends the use of Korean for the rest of the activity and the conversation continues entirely in English until the researcher later approaches to check on their progress (as shown in Excerpt (3)).

Excerpt (2): Abel and Jae follow the movements of a snail assigned to them. They measure the length of the body, how long it travels, etc.

1 Jae: SHELL      *nun mos mek-ci /*  
         shell      TOP    cannot eat-right

- (You can't eat the shell, can you?)
- 2            *ike-n*            *peleya-toy/*  
 this-TOP        discard-must  
 (You should throw this out)
- 3            *ike*    *man*    *mek-kol*  
 this    only    eat-and  
 (and eat only this)
- 4    Abel:    *e/*  
 yeah  
 (Yeah.)
- 5    Jae:        WE NEED TO COOK IT/  
 6                PUT THIS RIGHT KID=/  
 7    Abel:       =AND PUT IN        *elum*    AND WE COULD EAT IT RIGHT=/  
                   and put in        ice    and we could    eat it right  
 (and if you put ice in it you could eat it, right?)
- 8    Jae:        =YEAH/  
 9                [(unintelligible)
- 10   Abel:        <sup>l</sup>*ike*    *nemwu*        *ttakttakhay-se*        *meke/*  
                   this    too                    hard-because        eat  
 (You eat this because it's too hard)
- 11   Jae:        *mos*            *meke/*  
                   Cannot        eat  
 (You can't eat it.)

- 12 Abel: *e ttakttakhay/*  
 yes hard  
 (Yeah, it's hard.)
- 13 Jae: [(unintelligible)]
- 14 Abel: [(unintelligible) *ha-myen cwuke/*  
 do-if die  
 (You die if (unintelligible).)]
- 15 Jae: (emphatically) SPEAK ENGLISH/
- 16 Abel: OKAY/
- 17 (touches the head of the snail) OOOOH/
- 18 Jae: NO LEAVE IT/  
 19 EY IT'S GONNA GO IN/
- 20 Abel: IF YOU SCARE HIM/  
 21 HE'S GONNA GO INTO THE SHELL RIGHT?/
- 22 Jae: ABEL JUST SEE/  
 23 (3.5) NOW WE DID MOUTH EYE FEET FEET FEET/  
 24 FEET FEET LEAVE HIM ALONE LIKE THAT ABEL/  
 25 IT'S GONNA GO IN/  
 26 SEE ITS FEET/
- 27 Abel: (3.5) (softly) IT'S GOING/

Excerpt (3) below grants some insight into Jae's preference for English as exemplified in (2) above. In (3), his fluent Korean conversation with the researcher

shows that his reluctance to speak Korean in (2) does not result from lack of proficiency, since his sentences are well-formed and apparently produced without difficulty. Rather, given the status of English both as the peer language of the young and as the language of the classroom, we may conclude that Jae's preference for English exemplified in (2) originates in his sense of English as the unmarked choice for classroom use (Myers-Scotton, 1993b). In addition, Jae's preference for English may have derived from his relationship with an older brother who first introduced Jae to speaking English. A child in an immigrant family is exposed to the second language in different ways depending on whether or not he/she is a firstborn child (see further, Shin, forthcoming). While firstborn children in immigrant families learn the second language when they enter school, younger siblings begin speaking the second language before they enter school through interaction with the older siblings at home. McClure (1981, p. 75), in her study of the children of Mexican immigrants in the Southwestern United States, notes that besides simply learning the second language earlier, younger siblings are often influenced by the language attitudes of the older siblings. Therefore, a child whose older siblings are well integrated into the mainstream community may identify more with the speakers of the dominant language and prefer to be associated more with that group. Interestingly, in the current bilingual corpus, all five children who produced almost no intra-sentential codeswitching (i.e. Jae, Abel, Joshua, Kyung and Matthew) are either second or third born children.

Excerpt (3): Snail observation continues. Researcher approaches Jae and Abel.

1 Res:     *eti     pwa     ta     hay-sse?/*

where see all do-PAST

(Let's see. Did you finish it?)

2 Jae: *ike twu-kay nun mos pwa-sse/*  
this two-CLASS TOP not see-PAST

(We didn't see these two.)

3 Res: *mwe etten twukay lul mos pwa-sse/*  
what which two ACC not see-PAST

(Which two didn't you see?)

4 Jae: *ike hako ike=/*  
this and this

(this and this)

5 Res: *=yoke lang SHELL hako FEELER hako mos hay-sse/*  
this and shell and feeler and not do-PAST

(You didn't see this and the shell and the feeler?)

6 *ike hako ike nun hay-sse/*  
this and this TOP do-PAST

(Did you do this and this?)

7 EYES *hako* FEET  
eyes and FEET

(eyes and feet?)

8 Jae: *ikes to hay-sse/*  
this also do-PAST

(I also did this)

- 9 Res: *ikes to hay-sse?/*  
 this also do-PAST  
 (You also did this?)
- 10 MOUTH *to hay-sse?/*  
 mouth also do-PAST  
 (You also did the mouth?)
- 11 *o kulaysse/*  
 oh that is  
 (Oh, is that so)
- 12 *Twulise kachi hay-sseABEL hako?/*  
 twosome together do-PAST Abel with  
 (Did you do it together with Abel?)
- 13 Jae: NAY/  
 (yes)

In the following section, I examine the mixed categories in further detail.

### **Types of intrasentential language mixing**

As Figure 2 shows, single words of English-origin in Korean contexts form the overwhelming majority of the data (73%)<sup>3</sup>, followed by English-origin multiword fragments (25%) and Korean-origin single words and multiword fragments in English contexts (2%). This result confirms findings in previous studies that single lexical items of one language origin inserted in the discourse of another comprise the greatest portion

of any bilingual corpus systematically studied (Poplack & Meechan, 1998). A single English-origin word in Korean context is a word that is bordered on both sides by Korean material, as in (6), or is bordered on only one side by Korean material, provided it appears in clause-initial or clause-final position, as in (7).

(6) David: **na** paper **kasko issukkey**.

I have be will

(I'll have the paper.)

(7) So Hee: pencil **hana pakkey epse**.

one only not have

(I only have one pencil.)

A single Korean-origin word in English context is a word that is bordered on both sides by English material, as in (8), or is bordered on only one side by English material, provided it appears in clause-initial or clause-final position, as in (9).

(8) Abel: And put in **elum** and we could eat it right?

ice

(And put in ice and we could eat it, right?)

(9) Gina: how much is **kkoskka**?

traditional Korean dress

(How much is the dress?)

In comparison, English-origin multiword fragments are constructions containing two or more consecutive words in English and are found in an otherwise Korean clause, as in (10) and (11).

(10) Grace: tiny one      **twu hayyatoy?**  
also do must

(Do we have to do the tiny ones, too?)

(11) Yooni: **wuli ka many story manduleyatoy.**  
we NOM make must

(We must make many stories.)

Finally, Korean multiword fragments are constructions of two or more consecutive words embedded in an otherwise English sentence, as in (12).

(12) Kathy: **mwusun sayk** I could use.  
what color

(What color I could use.)

*(Insert Figure 2 about here.)*

In the remainder of this article, I focus specifically on single words of English-origin in Korean contexts, the largest and also the most contentious category of mixed items in the ongoing debate on linguistic properties of codeswitching.

### **Single English-origin words: Categories**

I first examine which grammatical categories are most prone to be integrated into Korean. It has been hypothesized that the more highly bound the morphemes (such as inflections, case endings, function words), the less likely it will be to undergo borrowing. Hierarchies of ‘borrowability’ (e.g., Haugen, 1950; Muysken, 1980; Poplack et al., 1988;

Van Hout & Muysken, 1996) suggest that common nouns are by far the most frequently borrowed elements, followed by grammatical categories with progressively less lexical content and more morphosyntactic elements. Such hierarchies, sometimes based on the proportion each part of speech represents in a corpus of borrowed items, may simply reflect the distribution of particular grammatical categories in native-language materials, rather than their specific propensity to be borrowed into another language. If, however, the proportions of grammatical categories of borrowed items were compared with these proportions in the monolingual discourse, we would have a more valid indication of the tendencies of each category to accept other-language material.

Figure 3 shows cross-category comparison of the distribution of parts of speech of English-origin words in Korean contexts with the distribution of native Korean words in monolingual Korean contexts, which resulted when some 1,600 words of running text of monolingual Korean spoken by the Korean-English bilingual children were parsed. A survey of the English-origin material in Korean contexts in the current corpus shows, in accord with previous studies, that the overwhelming majority (69.3%) falls into the category of nouns, followed by verbs (12.4%), numerals (9.7%) and color terms (5.9%). Lone English adjectives, adverbs, and conjunctions each constitute less than 2% of the lone English words in Korean contexts. (13) through (19) are examples in the current bilingual corpus of lone English-origin lexical items found in Korean contexts according to category:

(13) Kwon:	robber	<b>i-nikka</b>	<b>ilekhey</b>	<b>kul-ays-cyanha.</b>
(noun)		COP-because	like this	do-PAST-not

(You did this because you're a robber, didn't you?)

(14) Yooni: **wuli kathi** put **hay-ca.**

(verb) we together do-let's

(Let's put it together.)

(15) Grace: **kuntay ike** five **i-cyanha.**

(numeral) well this COP-not

(Well, this is five, isn't it?)

(16) Gina: **ne** red **h-ayss-cyanha.**

(color) you do-PAST-not

(You did the red, didn't you?)

(17) Grace: **wuli** late **ha-myen an-toy.**

(adjective) we do-if not-work

(We must not be late.)

(18) Kathy: **nay ka** first **ta an h-aysse.**

(adverb) I NOM all not do-PAST

(I didn't do all of it first.)

(19) Kwon: because **nemwu ssaykay no-nikka.**

(conjunction) too hard put because

(Because you put it too hard.)

*(Insert Figure 3 about here.)*

It is clear from Figure 3 that the tendency for English nouns to be borrowed exceeds the frequency of native Korean nouns by a factor of about ten. Thus, we may conclude that English nouns have a particular propensity to be borrowed, regardless of the frequency of occurrence of nouns in Korean. In the next section, I employ a quantitative variationist method to determine the status of English-origin nouns in Korean discourse.

### **Quantitative analysis of English-origin nouns in Korean discourse**

An examination of the current bilingual corpus reveals that many English-origin nouns are either in direct object position (N=96) or in subject position (N=40) of Korean verbs in otherwise Korean discourse.<sup>4</sup> Recall that subjects in Korean are often dropped. In the current bilingual corpus, over half (N= 542) of Korean sentences occurred without an overt subject while a slightly smaller proportion (N=405) appeared with a native Korean subject. Recall also that Korean has flexible word order, partly due to the presence of postnominal particles. In the current bilingual corpus, however, all 96 English-origin direct objects preceded the Korean verb, thereby resulting in an (S)OV order. Similarly, all 40 English-origin subjects preceded the Korean verb.

Recall that the equivalence constraint on codeswitching prohibits codeswitching from occurring between any two sentence elements which are ordered differently in the two languages. Since object and verb are ordered differently in Korean and English (i.e. Korean -- SOV; English -- SVO), any English-origin direct objects of Korean verbs should have the properties of borrowings into Korean rather than of codeswitches, insofar as the equivalence constraint against codeswitching between verb and object is correct. In

order to determine the status of these contentious items, one can make use of the facts of variable case marking in Korean, as illustrated by examples (20) through (22) from the current bilingual corpus. Specifically, one can compare the rate of occurrence of overt accusative markers on English-origin direct objects with the rate of occurrence of overt accusative markers on native Korean direct objects. If English-origin lexical items are behaving as borrowings into Korean (i.e. showing syntactic integration), the actual rates of case marking would be similar for English-origin objects and native Korean objects. In contrast, a statistically significant difference between the distribution of accusative case marking for the two sets of objects would imply that English-origin items are behaving as codeswitches.

(20) Overt accusative marker: card **lul** **sale-wasseyo.**

ACC buy-came.

‘I came to buy a card.’

(21) No accusative marker: pencil  $\emptyset$  **kace-wa.**

get-come.

‘Bring a pencil.’

(22) Topic marker in the accusative position: shell **nun** **mos** **mek-ci.**

TOP not eat-right

‘As for the shell, you can’t eat it, right?’

The same reasoning can be applied to English-origin lexical items in subject position of Korean verbs. While a statistically significant correlation between the distribution of nominative case marking of English-origin subjects and that of native

Korean subjects would indicate that the English-origin subjects are behaving as borrowings into Korean, a statistically significant difference between the two sets of subjects would suggest that English-origin subjects are behaving as codeswitches. The equivalence constraint makes a different prediction in the case of English-origin subjects, however. Since the boundary between subject and verb is an equivalence site (i.e. a permissible switch site) in Korean-English bilingual situation, at least some of the English-origin items in subject position of Korean verbs are expected to be *bona fide* switches.

### **Procedure**

To test these predictions, English-origin nouns governed by Korean verbs were extracted from the current corpus. For controls, native Korean nouns in direct object position of Korean verbs were also extracted so that the young bilingual speakers' own vernacular monolingual speech can be compared to their bilingual material. The extracted English-origin objects were divided into three categories: (1) Acc. -- with an overt accusative marker, (2) Zero -- without an accusative marker, and (3) Other -- with a topic marker or a delimiter. Recall that delimiters in Korean carry quantificational information and do not make the structural distinctions typically associated with case. In the current analysis, the 'Other' category represents a combination of the topic marker *(n)un* and delimiters *na* 'just', *to* 'also', and *man* 'only' which occur in the accusative position in lieu of the accusative marker. The above procedure was repeated with English-origin nouns in subject position of Korean verbs so that variable nominative marking can be analyzed.

## **Results**

### ***Accusative***

As can be seen in Table 3, there is a remarkable congruence in the distribution of the two sets of objects. Given the similar distribution of English-origin and native Korean direct objects in the three categories of accusative case marking, we may conclude that most, if not all, of the English-origin nouns in direct object position are borrowings.

Another piece of evidence which supports this conclusion is that all of the English-origin direct objects are exclusively content words with none of them being pronominal while 32% of the native Korean direct objects are pronouns. The lack of English function words such as pronouns is exactly what we would expect from borrowed vocabulary (Poplack et al., 1988). Similar findings were reported by Sankoff *et al.* (1990, p. 80) who found that English-origin pronouns and articles were missing in Tamil discourse.

*(Insert Table 3 about here.)*

### ***Nominative***

When we compare the two types of language items in subject position, there is, at first glance, somewhat of a disproportion, as can be seen in the first two rows of Table 4. However, we must note that a staggering 92% of the Korean subjects are pronominal, whereas none of the English ones are. When native Korean pronouns are removed, we find remarkable congruence with no statistically significant difference in the distribution of the two sets of subjects (rows 1 and 3). What this suggests is that the English-origin

subjects of Korean verbs behave like native Korean subjects and are thus best considered borrowings.

*(Insert Table 4 about here.)*

We can conclude then that most of the English-origin nouns appearing in subject or object positions in Korean discourse are borrowings, if only for the nonce. Although I have not made an attempt to distinguish between established loanwords and nonce borrowings in the current study, this bears no negative consequence on the current results since nonce borrowings and established borrowings pattern exactly alike in terms of syntactic integration into the host language (Poplack *et al.*, 1988).

## **Conclusion**

In this paper, I have made an attempt to provide a systematic account of the characteristics of intrasentential language mixing of a group of young bilingual Korean American children. We have seen that language mixing within the sentence boundary is generally rare when considered as a proportion of the entire corpus. We have also seen that there is significant inter-speaker variability in the amount and type of language mixing -- while children who mixed languages more often generally spoke more Korean, those who produced little or no language mixing spoke almost exclusively in English. Children who were more proficient in Korean naturally spoke more Korean, leading to more opportunities for them to codemix in a mostly English-medium instructional setting. We have also seen that language preference influences children's codemixing patterns.

As children with a preference for English speak more English despite their abilities in Korean, their interlocutors adapt their own language choice accordingly by using English themselves. Evidence concerning the development of codeswitching behavior shows that both knowledge of the interlocutor's linguistic capabilities and sensitivity to the norms for code choice upheld by their interlocutors are the major factors determining language choice in bilingual children (De Houwer, 1990, 1995; Shin & Milroy, 2000).

The fact that the data for this study were gathered in a mainstream English-speaking classroom may be crucial in helping us understand why some children exhibited a particular preference for English. There is often a tremendous amount of pressure for immigrant children to speak English in the school (Wong Fillmore, 1991). As one of society's main socializing instruments, the school plays a powerful role in exerting social control over its students. It endorses mainstream, middle-class values, and children who do not come to school with the kind of linguistic and cultural background supported in the schools are likely to experience conflict (Romaine, 1995, p. 242). Children are, therefore, motivated to speak English, while at the same time, motivated to discontinue using their primary languages. Given this, more language mixing may have been observed had the audio-recordings been made in another setting with a different set of interlocutors, such as at home with family members. In addition, there was some indication that children's language mixing patterns may differ according to birth order. Later born immigrant bilingual children have been found to be generally more English-dominant than firstborns because they tend to receive less direct parental speech input in Korean (Shin, forthcoming). In sum, these results render support to the argument that

sociolinguistic factors and context are paramount in finding explanations for children's language mixing behavior (De Houwer, 1995).

The current study also found that single English-origin words, particularly nouns, in Korean contexts form the overwhelming majority of the data, supporting results in other studies on childhood and adult bilingualism. These items are particularly problematic in a description of the language mixing of the Korean-English bilingual children, because many of them lack indications of syntactic integration into Korean. Given the facts of the grammars of Korean and English, we have tested a consistent framework for distinguishing codeswitching from borrowing, making use of quantitative variationist methods. We have seen through the variable analysis of case marking that most of the English-origin objects and subjects of Korean verbs are best treated as borrowings rather than codeswitches. This analysis lends credence to the theory that single lexical items of other language origin take the shape of native language material when placed in non-equivalence sites.

Overall, it appears that nonce borrowing is a particularly productive process in young Korean-English bilingual children's speech. This is in accord with the observation made by Poplack, Wheeler, & Westwood (1989) that nonce borrowing is especially prevalent in typologically different language pairs than typologically similar language pairs. While similarities between the languages (for example, between Spanish and English) allow many permissible codeswitch boundaries, differences between the languages eliminate some permissible switch locations. Since some of the switch sites are lost due to syntactic differences in typologically different languages, items of the other language origin are syntactically integrated into the host language so as not to violate the

grammar of either language in contact. This result has special implications for the ongoing debate over the distinction between borrowing and codeswitching because it emphasizes the need for a clear distinction between the two language contact phenomena. Simply assuming all items of other language origin to be codeswitching often results in apparent violations of various proposed constraints on codeswitching, and does little to improve our understanding of these different language contact phenomena.

## **Acknowledgments**

I am grateful to Shana Poplack, Marjorie Meechan, and Lesley Milroy for their helpful feedback on earlier versions of this article. I also thank Catherine E. Snow and two anonymous *Applied Psycholinguistics* reviewers for their insightful and critical comments.

## Notes

---

<sup>1</sup> In this paper, I employ the Yale Romanization of Korean (Martin, 1992).

<sup>2</sup> This is a mainstream classroom with English as the language of instruction. It contrasts with a bilingual classroom whose focus is clearly on a bilingual medium.

<sup>3</sup> Proper nouns such as names of persons and places were not included in the count.

<sup>4</sup> No English-origin tokens were found to occupy dative slots governed by Korean verbs in the current bilingual corpus. The remaining six English-origin nouns in Korean contexts occupied locative and genitive slots respectively. Due to the small number of tokens in these slots, a quantitative analysis could not be performed.