Code Switching In Sequential Bilingualism: 
A Polish-English Case Study
Aleksandra Kasztalska
Purdue University
Abstract

This study examined the code switching of a six-year-old Polish learner of English. The speech data was originally collected by Krupa-Kwiatkowska (1997) and consisted of a series of transcripts, of which a set of 14 was chosen. The transcripts were first analyzed using the CLAN program (Computerized Language ANalYSIS Program), with special attention given to the learner’s mean length of utterance (MLU) and his production of long utterances during a period of 16 months. It was hypothesized that the child’s overall linguistic productivity would increase and that, with time, he would choose to interact with others in English more frequently, especially when surrounded by English-speaking monolinguals. In addition to using CLAN, the data was closely examined to determine whether the learner’s code switching could be a result of an over-application of Polish lexicon and grammar. The findings largely supported the hypotheses, suggesting that the child’s internal dominant language was likely Polish but that he frequently used code switching as a conversational strategy.
Code Switching In Sequential Bilingualism: A Polish-English Case Study

Once regarded as an obstacle to a child’s development, bilingualism has become the subject of much research in the last decades, partly due to a change in general attitudes toward bilingual instruction (Gollan & Ferreira, 2009; Heredia & Altarriba, 2001; Martin, Krishnamurthy, Bhardwaj & Charles, 2003). Significant evidence suggests that the benefits of raising a child bilingually may outnumber the costs and that early acquisition of two languages may stimulate the young speaker’s linguistic and cognitive growth (Hakuta & Hakuta, 1991).

One of the most studied phenomena related to bilingualism is code switching, or code mixing. This tendency to switch between two languages in a single conversation (Hoff, 2005) is characterized by the borrowing of words, phrases, and even morphological elements between languages (Martin et al., 2003). Code switching may seem inefficient as a cognitive strategy because it requires that an individual go back and forth between two lexico-grammatical systems. To some extent, studies support this claim: Sentence comprehension, for example, takes longer when the sentences contain words from more than one language, as opposed to one language (Heredia & Altarriba, 2001). Moreover, in picture-naming tasks, when bilingual participants are allowed to choose words from either language, naming takes longer if the bilinguals switch languages between subsequent pictures than if they use one language (Gollan & Ferreira, 2009).

Despite these apparent drawbacks, bilinguals code switch very frequently in natural and in controlled experimental settings, when given the choice to do so (Gollan & Ferreira, 2009). Gollan and Ferreira (2009) suggest that bilinguals either do not foresee the possible costs of code switching or that the practical benefits of code switching outnumber these costs, especially when naming delays associated with switching are measured in milliseconds. Because code switching...
follows observable rules and preserves the grammar of both languages, it can in fact enhance mutual comprehension between bilinguals (Hereida & Altarriba, 2001; Meisel, as cited in Hoff, 2005, p. 343).

Code switching occurs both in simultaneous bilingualism—the concurrent acquisition of two native languages – and in sequential bilingualism—the achievement of native-like fluency after another language has already been mastered (Hoff, 2005). Often, one of the languages will become dominant in the learner’s mind (Hereida & Altarriba, 2001). The reasons for choosing one language over another reflect a number of sociolinguistic factors, such as the relative statuses of the two languages or the frequency of their respective use. Rather than addressing these issues, the present study aims to examine in more detail how code switching emerges.

To better understand the mechanisms underlying code switching, speech data collected by Krupa-Kwiatkowska (1997) was analyzed. The data came from a six-year-old native speaker of Polish acquiring English. As in other cases of sequential, subtractive bilingualism, with increasing fluency in English the boy chose to interact with other bilinguals and monolinguals in the high-prestige language, or English, rather than in Polish (Hakuta & Hakuta, 1991). It was hypothesized that an analysis of the child’s speech would reveal an over-application of Polish grammar rules in his English, suggesting that he was still in the early stages of bilingualism, as defined by Hereida and Altarriba (2001) and that his dominant language was still Polish. This hypothesis was largely confirmed, though the findings also suggested that the child was capable of separating the two linguistic systems enough to maximize his comprehension by others.
CODE SWITCHING IN SEQUENTIAL BILINGUALISM

Method

Participants

This research aims to shed more light on sequential bilingualism by using a corpus from the CHILDES database (MacWhinney, 2007). The series of 22 transcripts collected by Krupa-Kwiatkowska (1997) records the speech of the author’s son over two years. Marcin, a native Polish speaker who was six years and two months old (6;2) at the beginning of Krupa-Kwiatkowska’s (1997) study, had just arrived in the United States with his mother and had only sporadic English instruction prior to the move. He was recorded interacting with children aged five through seven, who came from three different socio-linguistic backgrounds: other English-Polish bilinguals, English monolinguals, and a Spanish monolingual. Of the 22 transcripts, I chose to look more closely at 14 that featured English-Polish bilinguals and English monolinguals, with the goal of focusing on the boy’s code switching between English and Polish over a period of 16 months (see Appendix B to learn more about the Polish language).

The Polish-English bilinguals, as defined by Krupa-Kwiatkowska (1997), included Basia, 7, Scott, 8 and Robert, 7; Sarah, 8 was an English monolingual. However, in the present study Scott was grouped with Sarah due to his extremely limited comprehension of Polish and his lack of Polish productive abilities. It is also worth noting that the Polish-English bilinguals, despite showing a preference for English, spoke Polish as well as or better than English (Krupa-Kwiatkowska, 1997). To sum up, the final classification of Marcin’s conversants was the following: Polish-English bilinguals (Basia, Robert) and English monolinguals (Sarah, Scott).

Marcin’s age will sometimes be denoted as a numerical value in the following format: Year;Month, e.g., 6;2 would refer to the boy being six years and two months old.
Lastly, this study focused primarily on Marcin’s English utterances, as a result of his increasing preference for English.

Materials and Procedures

First, the 14 selected transcripts were manually examined. The researcher noted the boy’s speech patterns and irregularities, before using the CLAN (Computerized Language ANalysis) program to identify further tendencies that might not have been apparent during the initial investigation. CLAN allowed the researcher to study Marcin’s number of utterances that consisted of seven or more words (longest utterances), his mean length of utterance (MLU), and his hypothesized shifting preference from Polish to English during the 16-month period. To determine whether Marcin’s language preferences were changing, longest utterances were further divided into those which were in English and those which were in Polish. The data was examined in detail with the aim of identifying other trends in Marcin’s use of English and Polish, e.g., lexical substitutions and omissions.

The reason for choosing seven words as the threshold in the first CLAN analysis was that the child’s longest utterance in the first transcript consisted of exactly seven words: pseciez mama spadnie do sasiadow i co? (but mommy will fall down to the neighors and what then?) (Krupa-Kwiatkowska, 1997). It was hypothesized that the frequency of Marcin’s longest utterances would reflect his overall linguistic development. To further explore this hypothesis, the researcher also examined the relationship between Marcin’s overall number of utterances and his longest utterances over time. The MLU, another likely indicator of the child’s development, was predicted to significantly increase during the 16-month period. Moreover, lexical and morphological intrusions of Polish into Marcin’s English utterances were studied. These
included lexical substitutions, incorrect or redundant prepositions, problems with direct and indirect objects, and issues related to definite and indefinite articles.

It must be mentioned that nonsensical phrases and exclamations such as “oh oh oh!” (Krupa-Kwiatkowska, 1997) were ignored in the analysis of MLU and longest utterances, since these highly expressive phrases offered no insight into the child’s linguistic development. Finally, phrases containing code switching, like potem na nich leza na chicken poc(s) (then the chicken pox is on them) (Krupa-Kwiatkowska, 1997) were classified as either Polish or English to avoid creating an in-between group. Since these lexical intrusions were limited to single words or compounds, such mixed utterances were grouped with either Polish or English utterances based on the language which appeared to predominate in them.

Results

Overall linguistic development

Marcin’s overall MLU significantly increased during the 16-month period, from the MLU of 1.6 observed at the beginning of the study to an MLU of 3.70 recorded in the last transcript (with a high of 4.00, 11 months into the study; see Figure 1). These findings contrast with Miller and Chapman’s (1981) study, which outlined several stages of MLU expected for normally developing children (cf. Hoff, 2005). Although monolingual speakers usually reach an MLU of 4.00-4.49 around the age of 37-52 months, Marcin’s only reported MLU of 4.00 occurred when he was 88 months old (7;4) and his MLU tended to cluster below 4.00 afterwards. As expected, Marcin’s increasing linguistic productivity was reflected in his longest utterances. As Table 1 shows, the frequency of utterances with seven words or more increased during the 16-month period. With just one seven-word utterance observed at the beginning of the study, in the last

A phrase used as an adjective that precedes a noun, e.g. “in-between,” should be hyphenated.

The Results section summarizes the data collected during the experiment. You do not need to list all findings in the text; tables and figures are often more efficient, especially when a lot of numerical data was collected.

It is helpful to remind your reader what your variables, e.g. Longest Utterances, stand for.

“cf.” is short for “compare” or “consult,” and is used to contrast or compare like things. For other common APA abbreviations, cf. http://blog.apastyle.org/files/ap-a-latin-abbreviations-table-2.pdf.

Explain why certain data may have been excluded in the final analysis, e.g. why nonsensical phrases were ignored.
transcript Marcin used 39 utterances with at least seven words in each of them, including a high of 48 utterances 13 months into the study, at 7;4.

Because the observed change could have resulted from an overall increase in the length of recorded conversations, an additional analysis was performed. First, the number of overall utterances in each transcript was calculated; next, the ratio of longest utterances over the overall number of utterances was calculated for each transcript (see Table 1). The results reveal that during the 16-month interval the frequency of longest utterances did increase, from 0.50% of all utterances in a transcript (at 6;3) to 18.20% (at 7;7), with a high of 19.60% (at 7;4). The data thus supports the initial hypothesis, in that Marcin’s utterances did become longer throughout the duration of the original study.

Shift at around 7;1

Marcin’s MLU and longest utterances suggested that the data could be further divided into those collected before and after the boy was about seven years and one month old. Most of the transcripts collected before 7;1 recorded Marcin’s interactions with English monolinguals, and as can be seen in Figure 1 his MLU experienced a sharp increase during this period. In contrast, the boy’s MLU in interactions with Polish-English bilinguals between 7;1 and 7;7 did not increase by nearly as much. However, statistical tests would need to be conducted to confirm that the observed shift was indeed significant.

Marcin’s overall MLU seems to stabilize after 7;1, but it is not clear whether this stabilization was an effect of a general linguistic trend occurring at that time or if the shift resulted from growing interactions with bilingual speakers. More data before 7;1 is needed to...
determine if Marcin’s MLU in interactions with bilinguals underwent a boost before it stabilized or if the boost was limited to his interactions with English monolinguals.

Further hints of the abovementioned shift around 7;1 was found during an analysis of Marcin’s utterances with seven words or more. These data too could be divided into utterances recorded before and after the boy was seven years and one month old, at which time his longest utterances comprised 16.20% of all of his utterances. During the seven-month period before 7;1, Marcin’s longest utterances increased from 0.50 at 6;3 to 11.10 at 7;0 (constituting an increase of 10.60). In contrast, during the six-month period after 7;1 the number of longest utterances ranged between 11.50 and 18.20 (an increase of 6.70), with a high of 19.60. Just like Marcin’s MLU, it appears that the frequency of longest utterances increased until the boy was seven years and one month old, at which point the rate seems to slow down.

Preference for English and over-application of Polish

To determine whether Marcin’s language preference was shifting to English, longest utterances were examined in more detail. The data shows that in the first months of the study virtually all of Marcin’s longest utterances were in Polish, but that by the end of the 16-month period his longest utterances were almost entirely in English, even when the boy interacted with other bilinguals (see Figure 2). In fact, around 6;8 Marcin began communicating with other speakers almost exclusively in English.

As the analysis further revealed, Marcin’s increasingly complex English utterances were filled with lexical and morphological intrusions of Polish. A recurring example of this code switching was Marcin’s tendency to insert a Polish word into an utterance that otherwise contained English words, resulting in phrases like you gotta buy (. ) wozek (you gotta buy a cart).
or put w microwave (put in a microwave) (Krupa-Kwiatkowska, 1997). Although word-mixing did not appear to depend on context, the substitution of an English word or phrase with its Polish equivalent often occurred in phrases that called for uncommon or rather technical words, which a child of Marcin’s age may not be familiar with. An example can be found in Marcin’s suggestion, trzeba przedłużacz (we need an extension cord), which the boy uttered in the middle of an English conversation (Krupa-Kwiatkowska, 1997).

Lexical substitutions seemed to mostly, but not exclusively, affect open-class words. Prepositions were especially problematic, as Marcin tended to either insert a preposition where English would do without one (e.g., go in the downstairs), or to directly translate a Polish preposition into English (e.g., on picture) (Krupa-Kwiatkowska, 1997). Interestingly, although Marcin made most of such substitutions while speaking English, on a few occasions he chose an incorrect preposition while speaking Polish, as in the sentence to wytne () z nozyczkami, or I will cut it with scissors (Krupa-Kwiatkowska, 1997). The Polish sentence contains a redundant particle z, which in English is obligatory but in Polish is unnecessary because the noun nozyczki (scissors) is already inflected for the instrumental case.

In addition to word mixing, Marcin seemed to have trouble with English direct and indirect objects, especially in earlier transcripts, which often contained sentences like give me to me (Krupa-Kwiatkowska, 1997). Moreover, likely resulting from the lack of articles in Polish, Marcin tended to omit definite and indefinite articles in English, e.g., line is busy right now, or I smell camera (Krupa-Kwiatkowska, 1997). On several different occasions, the boy also mixed up the relative pronoun that with the interrogative what, as in: this is the part what we are doing now (Krupa-Kwiatkowska, 1997). This too appears to be motivated by an over-application of Polish, where the two kinds of pronouns often take the same form, co.
Discussion

As expected, the analysis of Krupa-Kwiatkowska’s (1997) corpus revealed that Marcin’s overall productivity increased during the original study period. The boy’s MLU increased from 1.60 at the beginning of the study to 3.70 in the last transcript, and although he initially used only a single seven-word utterance, he formed 39 utterances with at least seven words in the last transcript. Moreover, Marcin’s longest utterances consisted less than 0.50% of his total number utterances in the first transcript, but in later transcripts comprised almost 20% of his utterances. Further examination of the data also hinted at a stabilization of Marcin’s MLU around 7;4. In other words, even though Marcin was producing more utterances with at least seven words after 7;1, his mean length of utterance remained relatively constant. A possible explanation of this finding is that the boy was producing more utterances overall—both long and short—and that this general increase affected the longest Utterance scores but not the MLU averages.

Marcin’s MLU showed the greatest increase before 7;1, when the boy was interacting predominantly with English monolinguals. On the other hand, production of longest utterances remained more or less stable throughout the 16-month period. However, even though after 7;1 Marcin mostly interacted with bilinguals, his highest MLU score and longest Utterance score (4.00 and 48, respectively) came from a conversation with an English monolingual. The data was insufficient to make a definitive statement about the effects of the boy’s conversational partners on his productivity, but the results nevertheless hint at a possible advantage gained from interactions with English monolinguals.

Marcin’s speech appeared to undergo a change around the time the boy was seven years and one month old, with both MLU and longest utterances stabilizing around that point. Moreover, Marcin’s longest utterances showed an increasing preference for English, and in the
last transcript the boy was interacting with both bilinguals and English monolinguals almost exclusively in English. Most likely, with continuing exposure to English Marcin became more confident in the language. The data also showed that, with time, Marcin’s language of preference changed from Polish to English. This is consistent with previous findings on sequential, subtractive bilingualism, which suggest that a minority or low-status language tends to give way to the language regarded as high-status and/or used by the majority of speakers in a given setting (Hakuta & Hakuta, 1991).

Nonetheless, it is unlikely that this overt change in linguistic preference was related to a change in the child’s internal dominant language, as is often the case. Many bilinguals experience such shift after prolonged exposure to a second language, which results in recurring lexical and morphological intrusions of elements from their second language into their first language (Hereida & Altarriba, 2001; Martin et al., 2003). An examination of Marcin’s speech, on the other hand, revealed just a handful of English intrusions into his Polish utterances – in contrast with his over-application of Polish grammar rules in his English. For example, the boy used the English interrogative pronoun what instead of the demonstrative that in the phrase this is the part what we are doing now (Krupa-Kwiatkowska, 1997); this was likely due to the fact that a Polish speaker would use the same word co as an interrogative what? or a pronoun opening a dependent clause. This suggests that the boy was in the early stages of bilingualism, as defined by Hereida and Altarriba (2001), and his dominant language was still Polish.

The results of the study suggest that despite code switching’s apparent drawbacks (e.g., production delay) bilinguals engage in this activity quite frequently (Gollan & Ferreira, 2009; Meisel as cited in Hoff, 2005, p. 343). The present findings were also consistent with a Spanish-English bilingual study by Lindholm and Padilla (1978), who concluded that single-word
substitutions were much more frequent than whole-phrase substitutions (as cited in Hakuta & Hakuta, 1991, p. 147). Furthermore, a large number of English words and phrases that were often substituted with their Polish equivalents were uncommon or rather technical, providing support for the claim that bilinguals use their nondominant language on easy tasks but and dominant language on more difficult tasks (Gollan & Ferreira, 2009).

Marcin also tended to omit definite and indefinite articles (which Polish lacks) and to ignore direct and indirect pronouns. As a pro-drop language, Polish can do without subject pronouns, but direct and indirect objects cannot be omitted, and so it is unclear why Marcin would drop the latter. Despite numerous omissions, substitutions and flat-out errors, Marcin had little trouble making himself understood and rarely made grave grammatical errors in English. This finding further suggests that code switching is not really arbitrary, as it does not usually impede mutual comprehension between speakers (Hereida & Altarriba, 2001) and as the speakers are capable of separating the two linguistic systems in their minds (Martin et al., 2003; Hoff, 2005).

The practical benefits of code switching thus seem to outnumber the costs (Gollan & Ferreira, 2009). Perhaps, as Hakuta and Hakuta (1991) suggested, code switching is actually a reflection of the speaker’s advanced grasp of both languages. At the same time, it is important to remember that the road to fluency is long, and that during the original study Marcin was only beginning to switch to English as his preferred language, which may explain why the boy’s MLU was below average (Miller & Chapman, 1981). With ongoing exposure to English, the boy may have eventually achieved native-like comprehension and productivity, though he likely continued to code switch, as bilinguals code switch even more with age (Gollan & Ferreira, 2009).
References


## Appendix A

**Table 1**

*Increase in Marcin’s production of longest utterances relative to all utterances*

<table>
<thead>
<tr>
<th>Age</th>
<th>Longest utterances</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>6;3</td>
<td>1</td>
<td>0.50%</td>
</tr>
<tr>
<td>6;3</td>
<td>2</td>
<td>2.60%</td>
</tr>
<tr>
<td>6;3</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>6;7</td>
<td>13</td>
<td>7.20%</td>
</tr>
<tr>
<td>6;8</td>
<td>4</td>
<td>9.30%</td>
</tr>
<tr>
<td>6;9</td>
<td>12</td>
<td>8.50%</td>
</tr>
<tr>
<td>6;9</td>
<td>17</td>
<td>7.30%</td>
</tr>
<tr>
<td>7;0</td>
<td>24</td>
<td>11.10%</td>
</tr>
<tr>
<td>7;1</td>
<td>44</td>
<td>16.20%</td>
</tr>
<tr>
<td>7;3</td>
<td>14</td>
<td>11.50%</td>
</tr>
<tr>
<td>7;4</td>
<td>48</td>
<td>19.60%</td>
</tr>
<tr>
<td>7;4</td>
<td>45</td>
<td>17.00%</td>
</tr>
<tr>
<td>7;5</td>
<td>28</td>
<td>12.20%</td>
</tr>
<tr>
<td>7;7</td>
<td>39</td>
<td>18.20%</td>
</tr>
</tbody>
</table>

The table must have a clear and concise title. Headings should be clear and brief; they should not be much wider than the widest entry in the column. For more information on formatting Tables, see: [http://owl.english.purdue.edu/owl/resource/560/19/](http://owl.english.purdue.edu/owl/resource/560/19/).

Visual materials can quickly and efficiently present a large amount of information. However, only more complex data should be presented in tabular format, and data that would require only two or fewer columns and rows should be presented in the text.
Figure 1. Marcin’s MLU throughout the 16-month study period. “EN” refers to his interactions with English monolinguals and “BL” to interactions with Polish-English bilinguals.
Figure 2. Shift in Marcin’s preference from Polish to English in his longest utterances. “EN” refers to his interactions with English monolinguals and “BL” to interactions with Polish-English bilinguals.
Appendix B

An Indo-European language, Polish belongs to the West-Slavic family and, as such, is most closely related to Czech and Slovak. In general, Polish follows the Subject-Verb-Object word order also found in English, but as a synthetic language it has an extremely flexible syntax. This flexibility translates into a high movability of words on the clause and sentence level, and is made possible by the high inflectionality of Polish lexemes, which contain much more syntactic information than in English. For example, a verb may be morphologically to indicate when an action/event occurred (tense), whether the action/event is completed (aspect), who is the agent/actor (which is why Polish allows the omission of personal pronouns). The inflection can also mark the speaker’s mood and sometimes even gender.

At the same time, Polish verbs are nowhere nearly as complex as in English, Spanish or French—all of which have compound tenses like future perfect or past perfect that do not exist in Polish. This difference often leads to the difficulties that many Polish speakers have in the acquisition of the English tense system.

In comparison with the rather impoverished case system of contemporary, Polish are declensions are quite complex. Nouns generally have seven grammatical cases: nominative, genitive, dative, accusative, instrumental, locative, and vocative. Moreover, each Polish noun is assigned grammatical gender: masculine, feminine, or neuter. On the other hand, Polish has no articles, unlike many Romance and Germanic languages. Finally, it should be noted that Polish has no distinct word order to construct questions, which are simply marked by rising intonation. An exception is the interrogative particle czy; while it can be used in Yes/No Questions, the particle is not obligatory and is often omitted.