CHAPTER 5

Triggered code switching

Evidence from Dutch – English and Russian – English bilinguals

Mirjam Broersma, Ludmila Isurin, Sybrine Bultena, and Kees de Bot
Radboud University Nijmegen/ Ohio State University/Radboud University Nijmegen/ University of Groningen

This paper presents further evidence that cognates may facilitate code switching. In three corpora of natural speech, code switches occur more often directly following or in the same clause as a cognate (or 'trigger word') than elsewhere. Triggered code switching is found between typologically similar languages, with Dutch L1 – English L2 speakers in New Zealand and Australia, and between typologically distant languages, with a Russian L1 – English L2 speaker in the USA. We find that words following but not words preceding a trigger word have an increased chance of being code switched, that form overlap without meaning overlap may be sufficient for triggering to take place, and that the attachment of extensive Russian morphology to a trigger word stem does not diminish its triggering potential. We do not find an effect of the interlocutor's use of trigger words. Further, discourse connectors are often used in the vicinity of code switches.

Keywords: code switching, cross-linguistic comparison, triggering, psycholinguistics

Introduction

In the mind of the multilingual, words, syntactic structures, phonology, and gestures from at least two languages wait for a chance to be produced. They all have their own language specific form and can be produced in perfectly monolingual utterances, but more interestingly, they can also be mixed in all possible combinations. Russian words can be produced with an English word order, English words can be embedded in Russian morphology, and these morphemes might in turn be

pronounced with an English–like phonology. There is no guarantee that a language choice on one of these levels is paralleled with a similar choice at another level. However, there is no guarantee either that a language choice on one level might not influence choices at other levels.

In De Bot et al. (this volume), we argue that when it comes to code switching, the language system of a bilingual in a code switching setting might be like a pile of sand. When more sand drips onto the pile, at some point the sand will start shifting and smaller or larger avalanches will take place. When and where the avalanches will occur and how many and how big they will be is unpredictable. Also, it is impossible to point to a single grain of sand and say that it 'caused' the avalanche, but still its contribution might have been crucial – just like that of all the other grains of sand. Likewise, a minor change in the language system may lead to a major shift, i.e., a change in the output language.

We assume that the different levels of speech might interact, and a shift from one language to another on one level (say, the word level) might stimulate a shift towards that language on other levels (e.g., syntax and phonology) as well, but again, this might be a matter of likelihood and by no means a rule. Similarly, of course, a shift on one level might have a continuing influence at that same level: when Russian becomes the 'selected' language at the syntactic level, it may stay selected for a while. This becomes interesting in those cases where the shift was an 'accidental' one. For example, if two languages share a certain syntactic construction, the use of that construction might stimulate a shift from one language towards the other language, even if the construction was not produced in that other language. Thus, the similarity between the two languages might in this case *trigger* a code switch.

In this paper, we investigate how a shift in language dominance on one level influences the next language choice *at the same level*. We focus on the lexical level, and we assess several issues that are important there: which words are likely to be code switched under the influence of a trigger, what is the role of word meaning and word form in triggering a code switch, what is the role of morphological complexity, can hearing a trigger word spoken by an interlocutor also trigger code switches, and can repetitive discourse elements trigger code switches?

In order to address these different questions, we look at different combinations of languages. First, we discuss code switching between two typologically similar languages, namely Dutch and English. We present data from two corpora of bilingual speech, with Dutch L1 – English L2 speakers living in New Zealand and Australia. Next, we discuss code switching between two typologically distant languages, with data from a Russian L1 – English L2 speaker living in the USA. The data from these different combinations of languages provide us with several new insights into triggered code switching. We apply statistical tests to natural

speech data, an approach that combines the benefits of naturalness with those of generalizability.

Of course, there are different reasons for people to code switch or not to code switch. The situation we are interested in here is one in which speakers feel free to code switch, and in which a small trigger might be enough to tip the balance, or rather, to make the sand slide and cause an avalanche.

Triggered code switching at the lexical level

So far, the existence of triggered code switching has only been established at the lexical level. In the 1960s, Michael Clyne already observed that code switches seemed to occur relatively often in the neighborhood of a cognate (Clyne, 1967). He proposed that cognates might *trigger* a switch from one language to another (Clyne, 1967, 2003). We elaborated on this idea and tested it in a corpus of natural, self-recorded Moroccan Arabic – Dutch speech (Broersma & De Bot, 2006). The results of statistical testing showed that, indeed, code switches occurred more often when a cognate was present. We found that 1) words that immediately followed a cognate were switched more often, and 2) in addition to that, words that did not directly follow a cognate but were part of the same basic clause as the cognate were also code switched more often than was the case in the absence of a cognate.

Our explanation is that the selection of a cognate leads to an increased activity of the 'other' language at the lexical level. As cognates are very similar in two languages, they are likely to be strongly connected at the lexical level. Thus, when a bilingual speaks Dutch, selection of a cognate might lead to an increased activation of Moroccan Arabic and vice versa, which might enhance the occurrence of code switching.

Dutch and Moroccan Arabic are typologically dissimilar and do not share many cognates. In the corpus, all cognates were nouns, many of them proper nouns. The proportion of cognates was relatively low, namely 5% (104 cognates on a total of 2224 words). It is possible that for such dissimilar languages, with so few cognates, the impact of a cognate is much larger than for languages that overlap in many ways. Therefore, we would like to investigate whether for languages that are typologically similar and share many cognates, these cognates also induce code switching. To this end, we selected two languages that are lexically strongly related: Dutch and English. In the two corpora collected among Dutch – English bilinguals in New Zealand and Australia, described below, 68% and 64% of all words were cognates (9336 and 5956 cognates on a total of 13648 and 9344 words, respectively), and cognates occurred in all grammatical classes.

Further, we wanted to investigate which words exactly can trigger a code switch. In Broersma and De Bot (2006), trigger words were defined as cognates, sharing both word form and word meaning across languages, allowing for small differences in phonological form. In the present paper, we investigate the role of meaning and form overlap in triggered code switching. Is word form overlap enough to trigger a code switch, or is word meaning overlap an essential requirement? Thus, do false friends trigger code switches in the same way as cognates do? The Moroccan Arabic – Dutch data used in Broersma and De Bot (2006) did not contain any words that overlapped in form but not in meaning. As Dutch and English share so many cognates, it turned out to be impossible to investigate this question with this language pair, as the data simply do not provide enough non-cognates to separate the role of form and meaning. Therefore, we turn to another pair of typologically dissimilar languages that do share enough false friends to investigate the role of form and meaning overlap separately: Russian and English. In this data set, 3% of all words were trigger words (86 trigger words on a total of 2896 words).

Elaborating on the question which words can trigger a code switch, we also investigate how transparent a cognate needs to be in order to function as a trigger. As Russian is a morphologically complex language with an extensive system of prefixation and suffixation, cognates are in many cases not easily recognizable as such. Cognate stems are often embedded in morphological structure, as in the example of *za-gaz-ovannom*, containing the cognate *gaz*. The question then becomes whether such morphologically masked cognates can still trigger a code switch.

Further exploring the Russian – English corpus, we investigate whether trigger words uttered by one speaker can induce code switching by another speaker. Finally, we address the use of discourse connectors that do not overlap in form and meaning in the two languages, but that have similar patterns of use in both languages. Are such discourse connectors related to code switching, and if so, are they cause or effect of the code switching?

Dutch – English data from New Zealand

Materials

First, we investigated the effect of cognates on code switching in typologically similar languages with strong lexical overlap. The materials used here were collected in a study into language loss among Dutch immigrants in New Zealand (Hulsen, 2000). Six interviews with Dutch – English bilinguals that seemed to contain a substantial amount of code switching were considered for analysis. After

Table 1. The informants (AoA: age of arrival in New Zealand; Length of residence: in New Zealand)

Informant #	Gender	Age	AoA	Length of residence	Interview duration (min.)	Clauses with CS (%)
1	F	73	39	34	24	35
2	F	59	12	47	20	17
3	M	67	21	46	27	16
4	M	69	29	40	13	26

transcription, the four interviews with the largest number of code switches were selected. They contained between 36 and 138 code switched words (on a total of 1248 to 3417 words), and between 34 and 106 clauses containing a code switch (on a total of 132 to 412 clauses), with a proportion of clauses containing a code switch between 16 and 35%. (For the two excluded interviews, these values were all smaller: 8 and 25, 7 and 28, and 4 and 15% respectively.)

The informants, two male and two female speakers, were between 59 and 73 years old at the time of the interview. They had moved from the Netherlands to New Zealand when they were between 12 and 39 years old, and had been living there between 34 and 47 years (Table 1).

During the interview, informants were asked about their experiences around their immigration, their life in New Zealand, visits to the Netherlands, and their use of and attitude towards the use of Dutch and English. Interviews were carried out in the informants' homes. The interviewer was a Dutch female, and whereas the respondents were aware that she knew English well, she spoke only Dutch during the interview. The interviewer's contribution to the conversations was limited to questions and short responses, and intended to elicit spontaneous, running speech from the informants. Only the informants' speech is examined here.

Method

Two main questions were addressed. First, are words directly following a trigger word code switched more often than other words? Second, are words that do not directly follow a trigger word but that are part of the same basic clause as the trigger word also code switched more often than other words? In Broersma and De Bot (2006), positive evidence was found for both questions. The former approach, based on the word order of the utterance, stays close to Clyne's original view on triggered code switching (Clyne, 1967, 2003), the latter, based on the clause level, stems from Broersma and De Bot's (2006) elaboration of triggered code switching. These two main questions were assessed as follows.

First, each word spoken by each informant was coded as a trigger word or a non-trigger word. Words were considered trigger words when they overlapped both in form and meaning across the two languages, allowing for some differences in the two languages. Trigger words thus comprised cognates and proper nouns. As lexical activation concerns the activation of lemmas, not lexemes (e.g., Levelt, 1989), whether a word was a trigger word was also determined on the basis of its lemma form. E.g., all forms of the verbs 'to go' / *'gaan'* were considered trigger words, including the partially overlapping lexemes 'go' / *'ga'* and the entirely non-overlapping lexemes 'went' / *'ging'*. Trigger words comprised translation equivalents that are fully homophonous with identical IPA notations, forms that are phonetically very close, and forms that are not homophonous in citation form, but very similar in running speech and in the informant's pronunciation.

One judge (the first author of this paper) coded each word as a trigger word or a non-trigger word. Statistical analyses were based on this judge's coding. As the definition of trigger words entails a certain degree of subjectivity, in order to assess the reliability of this first judge's coding, five other judges coded a subset of 21% of the data, deciding independently from one another for each word in the subset whether they considered it as a trigger word or not. All six judges were native speakers of Dutch who were proficient in English as a second language. The inter-rater reliability (calculated with a Two-Way Random Effects Model, Type Absolute Agreement, Average Measures Interclass Correlation Coefficient) was high (ICC =.914, p <.001), showing that there was strong agreement among all six judges. The correlation between the first judge's coding and each of the five other judges was also high (averaged Pearson r (2552) =.6248, p < 0.01). These results warrant the analyses based on the first judge's classification of trigger words and non-trigger words.

Next, for the word level analysis, the language of each word was established in a linear fashion. If a word differed in language from the preceding one, it was considered to be code switched. Note that trigger words are never counted as a code switch. Arguably, this leads to an underestimation of the amount of code switching. Such a rigorous categorization, however, is crucial for the current study, as it makes the statistical testing of the triggering theory possible. In the following examples, Dutch words are in italics, English words in non-italics, and trigger words in bold. For the intelligibility of the examples, trigger words are also spelled and italicized as if they are either Dutch or English; however, this is for clarity only and is not meant to suggest the trigger words actually belong to that language. Note also that the Dutch and English spelling of trigger words may differ considerably (e.g., 'you' / *'je'*), without indicating poor overlap in phonological form.

In the word level analysis, example 1 contains two code switched words ('chemist' and *'vroeg'*), both following a trigger word, and example 2 contains six

code switched words ('still', *'mensen'*, 'still', *'gordijnen'*, 'still', and *'nog'*), all following and sometimes also preceding a trigger word.

(1) Informant # 4:
Engels is zo **widely** understood **in Holland** / ***dat als je*** *zelfs naar* ***een*** *winkel* ***ging,*** *naar naar* ***een*** chemist shop ***of zoiets*** / ***en je*** *vroeg naar iets* / ***en het was in een beetje gebroken Hollands*** / ***dan*** *negen van* ***de tien*** *keer* ***het*** *winkelmeisje antwoordde* ***je*** *terug* ***in het Engels.***
[English is so widely understood in Holland that even if you went to a shop, to to a chemist shop or so, and you asked for something and it was a bit in broken Dutch, then nine out of ten times the salesgirl answered you back in English.]

(2) Informant # 3:
En dat vond ik het *leukst* / *om weer terug* ***te gaan*** / ***dat al die karakteristieken, die typische Hollandse karakteristieken*** **are** still there. ***De*** *mensen,* still ***de open*** *gordijnen,* ***en*** still **the** same / **what I** would call very ***typische Hollandse dingen*** / ***die zijn er*** *nog.*
[And that I liked best about going back, that all those characteristics, those typical Dutch characteristics are still there. The people, still the open curtains, and still the same what I would call very typical Dutch things, they are still there.]

Thus, a word was considered to be code switched when it was part of a different language than the previous non-trigger word. Note that, due to the high proportion of trigger words, the previous non-trigger word may be some words away. E.g., in example 1, *'vroeg'* is code switched relative to 'shop'. The four intervening words are all trigger words. Even the ones that have a clearly Dutch form might function as trigger words due to their overlap with English words, and therefore by definition they are not considered code switches here. Similarly, in example 2, *'nog'* is code switched relative to 'very', with six trigger words in between.

For the clause level analysis, the conversation was divided into basic clauses containing maximally one main verb. When a clause contained words from two languages, or from another language than the preceding clause, it was considered to contain a code switch. In the previous and in the following examples, basic clauses are indicated with a slash. In the clause level analysis, example 3 contains one code switch from the first to the second basic clause.

(3) Informant # 4:
Want ***er waren families hier*** / ***die dat deden*** on purpose.
[Because there were families here who did that on purpose.]

Example 4 contains two code switches, from the first to the second basic clause, and from the second to the fifth basic clause.

(4) Informant # 1:
En dan** praat **je Nederlands** /* because ***dat is much **more** */ **wat het was** / **niet wat het is** / maar **wat het was**.*
[And then you speak Dutch, because that is much more what it was, not what it is, but what it was.]

Example 5 contains a code switch within the second basic clause.

(5) Informant # 2:
Wij spreken soms Nederlands**, maar **niet** vaak. / Maar **emotioneel voor mij is het heel important.
[We sometimes speak Dutch, but not often. But emotionally for me it is very important.]

Example 6 contains code switches within the first and sixth basic clause.

(6) Informant # 3:
Maar anyway, **I could / I could / fill your whole tape** with **the** with **the passport question /** because **it it's a** very very **deep** issue / which **is totally** / which **is** *onbegrijpelijk / want **andere landen staan de dubbele nationaliteit toe**.*
[But anyway, I could I could fill your whole tape with the with the passport question because it it's a very very deep issue which is totally which is incomprehensible, because other countries allow the double nationality.]

The relationship between trigger words and code switches was investigated with two statistical tests, the χ^2 test for independence and Fisher's Exact test. Both tests assess whether two variables are independent of each other. Fisher's Exact test is more accurate than the χ^2 test if one of the cells contains a small value, or if the marginal is very uneven. The probabilities from both tests are reported as p for the χ^2 test and P for Fisher's Exact test.

Results

First, for the word level analysis, it was investigated whether words that followed a trigger word were code switched more often than words that did not border on a trigger word. Indeed, words that followed a trigger word were code switched 15.7% of the time, whereas words that did not border on a trigger word were code switched only 7.2% of the time (Table 2). This difference was statistically significant ($\chi^2 = 11.40$, $p < 0.001$, $P < 0.001$).

Table 2. Number of words that are code switched, number of words that are not code switched, and percentage of words that are code switched, split by words that follow a trigger word and words that do not border on a trigger word

Following a trigger word	Code switch		
	Yes	No	% Yes
Yes	82	439	15.7
No	19	245	7.2

Next, it was investigated whether words directly preceding a trigger word were also code switched more often than words that did not border on a trigger word. This was not the case, as Table 3 shows ($\chi^2 < 1$, $P > 0.2$). Similarly, for words that followed a trigger word, the chance of being code switched did not further increase when they also preceded a trigger word (Table 4; $\chi^2 = 0.00$, $P > 0.5$). Thus, for the word level analysis, trigger words only influenced the words directly following them, such that those words were code switched more often than the other words.

Table 3. Number of words that are code switched, number of words that are not code switched, and percentage of words that are code switched, split by words that precede a trigger word and words that do not border on a trigger word

Preceding a trigger word	Code switch		
	Yes	No	% Yes
Yes	48	470	9.3
No	19	245	7.2

Table 4. Number of words that are code switched, number of words that are not code switched, and percentage of words that are code switched, split by words that border on a trigger word on two sides and words that only follow a trigger word

Bordering on a trigger word on:	Code switch		
	Yes	No	% Yes
Two sides	166	889	15.7
One side (only following it)	82	439	15.7

At the clause level, basic clauses with a trigger word contained a code switch 23.2% of the time, while other clauses contained a code switch only 11.4% of the time (Table 5). This difference was significant ($\chi^2 = 5.2$, $p < 0.05$, $P < 0.05$).

Table 5. Number of basic clauses containing a code switch, number of basic clauses not containing a code switch, and percentage of basic clauses containing a code switch, split by basic clauses containing a trigger word and basic clauses not containing a trigger word

Trigger word	Code switch		
	Yes	No	% Yes
Yes	253	838	23.2
No	8	62	11.4

Similarly to the results of Broersma and De Bot (2006), the present results show that 1) words that immediately followed a cognate were code switched more often, and 2) words that were part of the same basic clause as the cognate were also code switched more often than was the case in the absence of a cognate. Thus, the influence of trigger words on code switching is not restricted to typologically distinct languages. Even for language pairs that share a very large number of cognates, these cognates exert a strong influence on other words and enhance the occurrence of code switching.

Dutch – English data from Australia

In order to further support the findings from the Dutch – English data collected in New Zealand, we endeavored to replicate the results with data from Dutch – English bilingual speakers in Australia.

Materials

The materials consisted of two interviews with Dutch – English bilinguals. They were Dutch immigrants, both females, who were 81 and 76 years old and had been living in Australia for 35 and 54 years respectively at the time of interview (Table 6). They were interviewed as part of a project on long term language attrition, and had previously been tested in 1971 and 1987 using the same procedures (De Bot & Clyne, 1994). The interviews contained 78 and 139 clauses with a code switch (on a total of 487 and 808 clauses), with a proportion of clauses containing a code switch of 16% and 17% respectively.

The interviews were carried out in a similar way as those described in Vection 3, with the exceptions that the interviewer was an Australian native speaker of English who spoke Dutch well, and that in addition to answering questions about their immigration, life in Australia, and their language use, the informants were also asked to describe one photograph in Dutch and one in English. The interviewer used only Dutch during the interview (as in Vection 3), except when the informant was asked to describe a photograph in English.

Table 6. The informants (AoA: age of arrival in Australia; Length of residence: in Australia)

Informant #	Gender	Age	AoA	Length of residence	Interview duration (min.)	Clauses with CS (%)
1	F	81	46	35	32	12
2	F	76	22	54	46	14

Method

The data were analyzed as in the previous section. Here, however, only analyses at the clause level were performed. We expect to provide more extensive data from this population in the near future.

Results

Clauses with a trigger word contained a code switch 17.0% of the time, and clauses without a trigger word never did. Correcting for the zero in one cell (Table 7), the difference was found to be significant ($\chi^2 = 5.94$, $p < 0.05$, $P = 0.05$).

Table 7. Number of basic clauses containing a code switch, number of basic clauses not containing a code switch, and percentage of basic clauses containing a code switch, split by basic clauses containing a trigger word and basic clauses not containing a trigger word

Trigger word	Code switch		
	Yes	No	% Yes
Yes	217	1062	17.0
No	0	16	0.0

Russian – English data from the United States

With the Dutch–English data, we have shown that cognates can trigger code switches, not only for typologically distinct languages but also for related languages with strong lexical overlap (i.e., sharing many cognates).

In this section, we turn to a typologically distinct language pair again. Russian and English represent linguistically and typologically distant families: Slavic and Germanic. Moreover, Russian is a language typologically distant both from English and Dutch. It shares fewer cognates with English, and, as a result, the data in this study revealed fewer triggers than the Dutch – English data.

With the Russian – English data, we address the question which words can trigger a code switch. First, we try to replicate the results found with the Dutch – English data. Then, we assess whether false friends can act as triggers as well, and whether morphological embedding reduces a cognate's ability to trigger a code switch. Next, we investigate whether the use of trigger words by the interlocutor can trigger code switching in the informant's speech, and we investigate the relation between discourse connectors and code switching.

Materials

The materials consisted of a conversation with one Russian English bilingual speaker. The informant was a 35 years old female who had moved from Kiev, Ukraine, where she was born and raised, to the United States when she was 21 years old and had resided there for 14 years at the time of the interview. Code switches occurred in 18% (110 out of a total of 613) of the informant's basic clauses.

The conversation was carried out largely similar as those described in Sections 3.1 and 4.1. It consisted, however, of free conversation without pre-determined topics. The conversation took place in the informant's home. The interviewer (the second author of this paper) was a female native speaker of Russian with a good command of English, whom the informant knew well. Although the informant was aware that the interviewer spoke both Russian and English well, the interviewer spoke only Russian throughout the interview. Further, the interviewer spoke as little as possible without corroding the naturalness of the conversation. The conversation lasted for 45 minutes and elicited 2896 words from the informant.

The informant was raised in a Russian–speaking family and had attended a Russian school. At the time, Kiev was predominantly Russian–speaking. She reported that currently English was her dominant language, both at home and at work. She considered herself a habitual code switcher, as she very frequently switched between Russian and English when speaking to Russian – English interlocutors. Below is a sample of the informant's code switching registered in the

present data. In the following examples, Russian words are in italics, English words in non-italics, and trigger words in bold.

(7) *A ja perezhivaju, potomu chto* **South America** it's so deep, maybe you need some shots, like for malaria. So *ja nemnozhko perezhivaju*.
[I worry because South America it's so deep, maybe you need some shots, like for malaria. So I worry a bit.]

Method

The data was coded and analyzed as in the previous sections. All words were categorized as trigger words or non-trigger words. In the word level analysis, for non-trigger words it was determined whether they immediately preceded or followed a trigger word. All non-trigger words were coded as code switched or not. In the following example, 'American', 'Ukrainian', 'passport' and 'Dominican Republic' are trigger words.

(8) No, he is an **American** citizen, which is going there weird, *potomu chto* I have to bring my **Ukrainian passport** *so mnoj, kogda my edem tuda v* **Dominican Republic**.
[No, he is an American citizen, which is going there weird, because I have to bring my Ukrainian passport with me when we go there, to the Dominican Republic.]

For the clause level analysis, each conversational sample was divided into basic clauses and each basic clause was then analyzed for the presence of trigger words and code switches. Example 9 illustrates how clauses were identified:

(9) And they go through **Italy** / and they went through such hard times / *kogda oni cherez* ***Italiju*** *ehali* / and had such a horrible time / to adjust here.
[And they go through Italy, and they went through such hard times when they go through Italy and had such a horrible time to adjust here.]

In the above example, the trigger word 'Italy' occurs twice, first in its English variant and second in Russian. The utterance begins in English, and is continued in Russian in the third basic clause. This basic clause is thus code switched compared to the previous basic clause, and also contains a trigger word, namely the word 'Italy' in its Russian variant.

Results

Cognates

Most of the trigger words in the present data were nouns. There were 86 trigger words on a total of 2896 words (3%). Note that this is less than for the Moroccan Arabic – Dutch data (Broersma & De Bot, 2006), where 5% of the words were trigger words.

The Moroccan Arabic – Dutch data as well as the Dutch – English data in Vection 3 of this paper show that words directly following a trigger word are more likely to be code switched than other words. The analysis of the Russian – English data showed similar results. Words directly following a trigger word were code switched 28.4% of the time, and other words only 9.5% of the time (Table 8). This difference was statistically significant ($\chi^2 = 54.2$, $p < 0.001$, $P < 0.001$).

Table 8. Number of words that are code switched, number of words that are not code switched, and percentage of words that are code switched, split by words that follow a trigger word and words that do not border on a trigger word

Following a trigger word	Code switch		
	Yes	No	% Yes
Yes	44	111	28.4
No	226	2146	9.5

Table 9. Number of words that are code switched, number of words that are not code switched, and percentage of words that are code switched, split by words that precede a trigger word and words that do not border on a trigger word

Preceding a trigger word	Code switch		
	Yes	No	% Yes
Yes	18	137	11.6
No	226	2146	9.5

Words preceding a trigger word were not code switched more often than words that did not border on a trigger word (Table 9, $\chi^2 < 1$, $P > 0.2$), nor was the chance of being code switched for words that followed a trigger word higher when they also preceded a trigger word (Table 10, $\chi^2 < 1$, $P > 0.4$).

Table 10. Number of words that are code switched, number of words that are not code switched, and percentage of words that are code switched, split by words that border on a trigger word on two sides and words that only follow a trigger word

Bordering on a trigger word on:	Code switch		
	Yes	No	% Yes
Two sides	2	8	20.0
One side (only following it)	44	111	28.4

At the clause level, basic clauses containing a trigger word contained a code switch 31.0% of the time, and other clauses only 13.9% of the time (Table 11). This effect of trigger words was statistically significant ($\chi^2 = 22.10$, $p < 0.001$, $P < 0.001$).

Table 11. Number of basic clauses containing a code switch, number of basic clauses not containing a code switch, and percentage of basic clauses containing a code switch, split by basic clauses containing a trigger word and basic clauses not containing a trigger word

Trigger word	Code switch		
	Yes	No	% Yes
Yes	45	100	31.0
No	65	403	13.9

Thus, the present data provide additional support to the findings from the Dutch – English data described above, as well as the Moroccan Arabic – Dutch data from Broersma and De Bot (2006). Code switches into another language are more likely to occur right after a trigger word, and basic clauses containing a trigger word are also more likely to contain a code switch.

Form and meaning overlap

The above analyses were based on triggers that were cognates, i.e., words that overlapped semantically and phonologically in two languages. Next we investigated whether word form overlap alone is enough to induce a code switch. The present set of data revealed numerous cases of false friends. False friends are words that overlap in their phonological form but differ in meaning in two languages. Costa et al. (2006) suggest that the only way in which false friends can be activated in the non-target language is through activation of the phonological form of the word in the target language that sends activation to the lexical form of the corresponding homophone in the non-target language. According to this argument, we might expect the Russian word *god* ('year') to activate the corresponding homophone

'god' in English and act as a trigger in activating the non-target language and subsequently inducing a switch into English. Some examples of false friends identified in the present data are given in Table 12.

Table 12. False friends

Word form in English / Russian	Meaning of the Russian word
dome / *dom*	house
family / *familia*	last name
god	year
hotel	wanted (3rd person masculine, past tense)
let	years (Genitive case)
on	he (Nominative case)
quarter / kvartira	apartment

Table 13. Number of words that are code switched, number of words that are not code switched, and percentage of words that are code switched, split by words that follow a false friend and words that do not border on a false friend or on a trigger word

Following a false friend	Code switch		
	Yes	No	% Yes
Yes	11	40	21.6
No	204	2016	9.2

In order to be able to assess the effect of false friends, words that bordered on a trigger word were removed from the analysis. Words directly preceding a false friend were not code switched more often than words that did not border on a false friend ($\chi^2 < 1$, $P > 0.3$).

As Table 13 shows, words directly following a false friend were code switched 21.6% of the time, and words that did not border on a false friend only 9.2% of the time. This difference was statistically significant ($\chi^2 = 8.91$, $p < 0.01$, $P < 0.01$). Thus, similarly to the trigger words, false friends did not affect code switching of the preceding words, but did lead to a higher chance of code switching of the following words. The words that followed a false friend (Table 13) were compared to the words that followed a trigger word (Table 8). There was no statistically significant difference in the percentage of code switching of these items ($\chi^2 < 1$, $P > 0.2$).

Thus, the results showed that false friends also triggered code switching, which suggests that word meaning overlap may not be necessary and word form overlap alone may be sufficient for triggering a code switch.

Morphological masking

Now that we have established that form overlap plays an important role in enhancing code switching, we turn to the question how much form overlap is needed, and specifically what the role of morphological complexity is. To what extent can bound morphemes mask a cognate that they are attached to and hinder its triggering potential? The morphological structure of Russian is very rich and provides the main source of word formation in the language. While there are cognates with an almost complete overlap in meaning and form, like *'transport'*, *'America'*, and '*mama*', some show only a partial resemblance with the form of their English counterparts due to prefixation and suffixation, e.g. *'tematicheskij'* where only the stem *'tema'* ('theme') retains its original form.

To investigate whether morphological complexity influences triggering, trigger words that were produced as Russian words were divided into three categories based on their morphological form: transparent, less transparent, and masked trigger words. The first category includes cognates with a zero ending or only one inflectional morpheme (e.g., *'moment'*, *'Kieve'*, *'sestra'*, *'mamu'*). The second includes cognates with a derivational suffix and words with both a derivational suffix and an inflectional morpheme (e.g., *'nacionalistka'*, *'radikalno'*, *'practicheski'*). The third includes words with both suffixes and prefixes, where only the stem retains the original meaning while Russian morphemes on both sides make it hard to recognize the word as an original borrowing from English (e.g., *'po-angliiski'*), and morphological blends (e.g., *'proftehshkola'*). An illustration of some of the trigger words found in the data and their classification are given in Table 14.

There were 77 instances of trigger words articulated in Russian. The majority (57) came from the first category (transparent), while the second category (less transparent) accounted for 15 trigger words, and the third (masked) for 5 trigger words.

Words that preceded a trigger word did not have an increased chance of being code switched for any of the three categories (transparent triggers: $\chi^2 < 1$, $P > 0.2$; less transparent triggers: $\chi^2 = 1.58$, $p > 0.2$, $P > 0.2$; masked triggers: $\chi^2 < 1$, $P > 0.3$).

The results for the words following a trigger word in each of the three categories are shown in Tables 15, 16, and 17. Words that did not follow a trigger word were code switched 9.5% of the time, words that followed a transparent trigger 17.5%, words that followed a less transparent trigger 20.0%, and words that followed a masked trigger 40.0% of the time. Comparisons of the three categories of triggers did not reveal any statistically significant differences (transparent versus less transparent triggers: $\chi^2 < 1$, $P > 0.5$; less transparent versus masked triggers: $\chi^2 < 1$, $P > 0.4$; transparent versus masked triggers: $\chi^2 = 1.5$, $p > 0.2$, $P > 0.2$). Thus, triggered code switching occurred regardless of the amount of morphological embedding.

Table 14. Morphological classification of Russian triggers

Transparent	Less transparent	Masked
dezajner ('designer')	adaptirovalas ('adapted')	po-angliiski ('in English')
institute	blondinka ('blond')	po-russki ('in Russian')
Kieve ('Kiev')	dochka ('daughter')	proftehshkola ('professional technical school')
lingvist ('linguist')	doktorskuju ('doctor')	
mama	immigracija ('immigration')	
moment	konstruktorom ('constructor')	
professoru ('professor')	nacionalistka ('nationalist')	
sestra ('sister')	practicheski ('practically')	
syn ('son')	radikalno ('radically')	
	tehnicheskoje ('technical')	

Splitting the trigger words into three categories considerably reduced the statistical power of the analyses. Indeed, the triggering effect did not reach significance for all categories of trigger words separately. For transparent triggers, the effect was significant ($\chi^2 = 4.1$, $p < 0.05$, $P < 0.05$), for less transparent triggers it was not ($\chi^2 = 1.9$, $p > 0.1$, $P > 0.1$), while for masked triggers, χ^2 yielded a significant result whereas Fisher's Exact Test just missed significance ($\chi^2 = 5.3$, $p < 0.02$, $P = 0.075$).

The finding that there were no statistically significant differences between the three categories of trigger words suggests that morphological embedding does not reduce the triggering capacity of a cognate. As even the combination of Russian prefixes and suffixes of the masked trigger words did not reduce the trigger words' power to trigger code switching, it seems that triggering is caused by the stem of the trigger word, regardless of its morphological structure.

Table 15. Number of words that are code switched, number of words that are not code switched, and percentage of words that are code switched, split by words that follow a transparent trigger word and words that do not border on a trigger word

Transparent trigger word	Code switch		
	Yes	No	% Yes
Yes	10	47	17.5
No	226	2146	9.5

Table 16. Number of words that are code switched, number of words that are not code switched, and percentage of words that are code switched, split by words that follow a less transparent trigger word and words that do not border on a trigger word

Less transparent trigger word	Code switch		
	Yes	No	% Yes
Yes	3	12	20.0
No	226	2146	9.5

Table 17. Number of words that are code switched, number of words that are not code switched, and percentage of words that are code switched, split by words that follow a masked trigger word and words that do not border on a trigger word

Masked trigger word	Code switch		
	Yes	No	% Yes
Yes	2	3	40.0
No	226	2146	9.5

Triggering between speakers

Recent studies on cross-linguistic priming show that what a speaker has just heard can affect his or her production (Hartsuiker, Pickering, & Veltkamp, 2004; Kootstra, Van Hell, & Dijkstra, this volume). If cross-linguistic priming is possible across the boundaries of perception and production, can we expect a triggering effect of the interlocutor's speech on the informant's speech?

So far we have analyzed the speaker's utterances without considering the possible effect of the interlocutor's utterances, as the latter were removed from the analysis. In the next analysis, the informant's speech is analyzed taking lexical triggers in the interviewer's speech into account. First, trigger words were identified in the interviewer's speech and next, code switches in the informant's speech after the interviewer's comments were identified. In order to identify a code switch, the language of the last basic clause in the informant's previous utterance was compared with the first basic clause in the informant's next utterance following the interviewer's turn. As mentioned before, the language of the interviewer remained the same, i.e., Russian. Consider the following discourse samples:

(10) Informant:
So I feel alone, *znaete, kogda moja mama uehala* and my father passed away.
[So I feel alone, you know, when my mom left and my father passed away.]
Interviewer:
Jeto vasha tetja kotoraja ***lingvist?***
[Is this your aunt who is a linguist?]
Informant:
Jeto s moej mamy storony, ona s Ukrainy.
[This is from my mother's side, she is from Ukraine.]

(11) Informant:
Odin god.
[One year.]
Interviewer:
A ***photographia*** *u vas est'?*
[Do you have a photo?]
Informant:
No, not in this purse. He is very handsome, tall, dark.

In examples 10 and 11, the interviewer's comment contained a trigger word (*'lingvist'* and *'photographia'*) and the language of the informant changed from English to Russian in example 10 and from Russian to English in example 11.

Table 18. Number of basic clauses containing a code switch, number of basic clauses not containing a code switch, and percentage of basic clauses containing a code switch, split by basic clauses preceded by a perceived trigger word and basic clauses not preceded by a perceived or produced trigger word

Perceived trigger word	Code switch		
	Yes	No	% Yes
Yes	12	21	36.4
No	22	23	48.9

There were 88 utterances from the interviewer in the data and almost half of these contained a trigger word. In order to be able to assess the effect of the interviewer's speech, basic clauses that contained a trigger word in the informant's speech were removed from the analysis, leaving 78 basic clauses for analysis. As Table 18 shows, when the interviewer produced a trigger word, the informant produced a code

switch 36.4% of the time, and when the interviewer did not produce a trigger word, the informant produced a code switch 48.9% of the time, which was not statistically different ($\chi^2 = 1.21$, $p > 0.2$, $P > 0.1$). Thus, there was no effect of the interviewer's use of trigger words on the informant's code switching behavior.

Note that overall, the percentage of code switches in Table 18 is relatively high compared to that in Table 11, namely 43.6% versus 17.9% (collapsing over both rows of each table). Table 11 represents all basic clauses uttered by the informant, and Table 18 only those that the informant uttered immediately after the interviewer had spoken. Thus, the informant code switched relatively often after the interviewer had spoken. This difference was statistically significant ($\chi^2 = 27.59$, $p < 0.001$, $P < 0.001$). As the interviewer always spoke Russian, we would expect the informant to switch from English to Russian more frequently than from Russian to English after the interviewer's utterances. Indeed, 20 of the informant's switches after the interviewer's utterances were from English to Russian and 14 were from Russian to English.

Repetitive discourse elements

So far, we have investigated the effect of a language shift due to the selection of trigger words that overlap in two languages in form and meaning, or in form only. However, it might also be possible that elements that do not overlap in either form or meaning may trigger code switches due to similarities in their patterns of use in the two languages. Here, we investigate whether there is a relation between the use of discourse elements and code switching. We propose that repetitive discourse elements are often used in the vicinity of code switches.

Thus, we are not looking for discourse elements that show semantic or phonological overlap in the two languages, but for frequently used patterns of discourse. The data revealed two such discourse connectors that were frequently used by the speaker: the phrase *vy znate* ('you know'), that was mostly used in its pro-drop version *znaete*, and a two-word conjunction *potomy chto* ('because'). The latter was mostly used as an intra-sentential connector, which legitimizes its classification as a discourse connector rather than a syntactic unit in the present analysis. The following examples illustrate the use of the two connectors (in bold print).

(12) *Ona zvonila, **znaete** kak, posylki tam, podarki*, and I grew up, ***znaete***, *jetot* Soviet Union fell apart and everything, you do not see any future over there, kind of scary and everything.
[She called, you know how it is, packages, gifts, and I grew up, you know, this Soviet Union fell apart and everything, you do not see any future over there, kind of scary and everything.]

(13) No, he is an American citizen, which is going there weird, ***potomu chto*** I have to bring my Ukrainian passport *so mnoj, kogda my edem tuda v* Dominican Republic.
[No, he is an American citizen, which is going there weird, because I have to bring my Ukrainian passport with me when we go there, to the Dominican Republic.]

(14) *Ja ego* renew ***potomu chto*** *u menja* Green Card and passport so I renewed it after a while. *No mne nado zanjat'sja uzhe* American citizenship, *pasport poluchit', **potomu chto*** I am weary.
[I renewed it because I have a Green Card and passport so I renewed it after a while. But I need to take care of the American citizenship, to get a passport, because I am weary.]

Example 12 shows two instances of *znaete* ('you know'). The first occurs in the Russian fragment, and the second connects an English fragment with a Russian word. Examples 13 and 14 illustrate the use of *potomu chto* ('because'). In 13, the Russian connector is embedded in an English fragment. In 14, it is preceded by English and followed by Russian on its first occurrence, and vice versa on its second occurrence.

There were 30 instances of *vy znate* and 47 instances of *potomy chto* in the data. Note that in nine cases the discourse connector itself was code switched, as it was used in an English context. However, in order to investigate the discourse connectors' relation to code switching, we consider them as borrowings, and thus as belonging to both languages, and we assess whether the words immediately before and after the discourse connector differ from each other in language. All cases where a trigger word might have played a role were excluded from the analyses.

Words preceding a discourse connector did not have an increased chance of being code switched ($\chi^2 < 1$, $P > 0.4$). Table 19 shows the results for the words following a discourse connector. Those words were code switched 38.5% of the time, while words that did not follow a discourse connector were code switched only 9.9% of the time. This difference was statistically significant ($\chi^2 = 54.2$, $p < 0.001$, $P < 0.001$).

Table 19. Number of words that are code switched, number of words that are not code switched, and percentage of words that are code switched, split by words that follow the discourse connectors 'vy znate' or 'potomy chto' and words that do not border on those or on a trigger word

'Vy znate' or 'potomy chto'	Code switch		
	Yes	No	% Yes
Yes	25	40	38.5
No	226	2063	9.9

The results show that *vy znate* and *potomy chto* preceded a code switch significantly more often than expected under chance. Note that this does not tell us anything about the directionality of the relation between those discourse connectors and code switching. We can think of several interpretations of this relationship. The first possibility is that the discourse connectors triggered the code switches. The similarity of the use of the discourse connectors in the two languages, or their frequent use in code switched utterances might have connected them strongly to both languages and given them triggering potential. The second possibility is that the code switches triggered the use of the discourse connectors. The speaker might have felt the need to embed her code switches in a particular structure, and thus the code switches might have called for the use of the discourse connectors. The third possibility is that there is one reason for both the use of the discourse connectors and the code switches. For example, if the speaker was faced with retrieval problems, she might resort to habitual discourse elements while trying to access the syntactic structure or lexical element in either of her two languages.

Thus, we do not wish to claim that the discourse connectors *vy znate* and *potomy chto* triggered code switching, but there clearly is some relationship between the two, such that the discourse connectors often occurred directly before a code switch.

General discussion

In this paper, we presented new evidence that the production of a cognate can trigger a code switch. In previous research (Broersma & De Bot, 2006), we showed that words spoken directly after a cognate or in the same basic clause as a cognate were significantly more often code switched than other words, in the speech of Moroccan Arabic – Dutch bilinguals. In the present paper, we replicate this finding with three data sets.

The amount of lexical overlap and the typological similarity of the languages in these data sets were widely different. In the Moroccan Arabic – Dutch data, 5% of the words were cognates, in the two Dutch – English corpora, this was 68% and 64%, and in the Russian – English data 3%. Typologically, both Moroccan Arabic and Dutch, and Russian and English are very different, whereas Dutch and English are typologically more similar. With all these language pairs, we found that cognates enhanced the occurrence of code switching. Thus, triggered codeswitching is not incidental and limited to particular language pairs but a robust and general phenomenon.

In line with the Moroccan Arabic – Dutch data, in the present data, code switches were more frequent in basic clauses containing a trigger word, as well as

immediately after a trigger word. There was no increase of code switching immediately before a trigger word.

We investigated which kinds of words can trigger a code switch. In the Moroccan Arabic – Dutch data, all trigger words were nouns, and most of them proper nouns. In the Dutch – English and Russian – English data, trigger words occurred in all grammatical classes. In all data sets, we found that cognates, overlapping in word form and in word meaning, triggered code switching.

Next, we investigated whether both form and meaning overlap were necessary for triggering to take place. With the Dutch – English data, the large number of cognates made it impossible to assess this question properly. With the Russian – English data, however, we assessed the effect of false friends, which overlap in word form but not in word meaning. The results showed that the presence of a false friend also led to an increased chance that the following word was code switched. Thus, it seems that form overlap alone may be sufficient to trigger a code switch.

We further investigated how transparent a cognate needs to be in order to trigger a code switch. We looked at Russian – English cognates that were, sometimes extensively, embedded in Russian morphology. The results showed that (while there was a triggering effect) there were no significant differences in the amount of code switching after transparent, less transparent, or masked primes, and that heavily masked cognates triggered code switches as much as transparent cognates without any morphological embedding did. Thus, triggering seems to be caused by the stem of a word, and even the extensive Russian morphology does not diminish a cognate's capacity to trigger a code switch.

With the Russian – English data, we investigated whether trigger words spoken by one speaker can trigger a code switch in the speech of another speaker. Although a host of sociolinguistic studies have shown the importance of the interlocutor's speech for code switching (e.g., Myers-Scotton, 1993), and recent research has shown that the interlocutor's code switching behavior has a direct impact on another speaker's code switching (Kootstra, Van Hell, & Dijkstra, this volume), we found that the interviewer's use of trigger words did not enhance code switching in the informant's speech. Thus, it seems that the trigger word needs to be produced rather than perceived in order to trigger a code switch.

Finally, we assessed the relation between discourse connectors and code switching. We found that code switches were very often preceded by discourse connectors. This might suggest that the use of discourse connectors induces code switching or vice versa, or both might result from the same cause.

In the present study, we found that both the word level analysis and the clause level analysis explained triggering well. We did not attempt to tease apart which analysis explained the code switching patterns in the data better. Previous work suggests that the clause level analysis explains code switching better than the word

level analysis when there are few trigger words and code switches in the data (Broersma & De Bot, 2006). The word level analysis, on the other hand, seems to give a better and more precise account of the data than the clause level analysis when the data contain a large amount of trigger words and code switches (Broersma, to appear). Those studies thus showed that both approaches have their merits, and that their predictions overlap to some extent, which the present research confirms.

The data were selected to contain a reasonable amount of code switching. Between 16 and 35% of all basic clauses contained a code switch in the speech of the four Dutch – English informants from New Zealand, and 12 and 14% did for the two Dutch – English informants from Australia, and 18% for the Russian – English informant. Thus, in these data, the informants code switched regularly, and their code switching was influenced by their production of cognates. We assume that when speakers do not feel free to code switch, it is not likely that a cognate will induce them to do so. Thus, the effects of cognates are presumably limited by social and pragmatic considerations (e.g., Blom & Gumperz, 1972; Myers-Scotton, 1993). Nevertheless, when the circumstances are such that the pile of sand is about to start sliding, a cognate might be the last grain of sand that causes the avalanche.

Acknowledgements

We are grateful to Madeleine Hulsen for making the New Zealand data available, to Michael Clyne for making the Australia data available, and to Don Winford and one anonymous reviewer for helpful comments on an earlier version of this paper.

References

Blom, J., & Gumperz, J.J. 1972. Social meaning in linguistic structures: Code-switching in Norway. In J. J. Gumperz & D. Hymes (eds.), *Directions in Sociolinguistics: The Ethnography of Communication*, 407–434. New York: Holt, Rinehart and Winston.

Broersma, M. (to appear). Triggered codeswitching between cognate languages. *Bilingualism: Language and Cognition*.

Broersma, M., & De Bot, K. 2006. Triggered codeswitching: A corpus-based evaluation of the original triggering hypothesis and a new alternative. *Bilingualism: Language and Cognition*, 9: 1–13.

Clyne, M. 1967. *Transference and Triggering: Observations on the Language Assimilation of Postwar German-Speaking Migrants in Australia*. The Hague: Martinus Nijhoff.

Clyne, M. 2003. *Dynamics of Language Contact: English and Immigrant Languages*. Cambridge: CUP.

Costa, A., La Heij, W., & Navarrete, E. 2006. The dynamics of bilingual lexical access. *Bilingualism: Language and Cognition*, 9, 137–151.

de Bot, K., Broersma, M., & Isurin, L. (this volume). Sources of triggering in code switching.
de Bot, K., & Clyne, M. 1994. A 16-year longitudinal study of language attrition in Dutch Migrants in Australia. *Journal of Multilingual and Multicultural Development*, 15: 17–28.
Hartsuiker, R. J., Pickering, M. J., & Veltkamp, E. 2004. Is syntax separate or shared between languages? *Psychological Science*, 15, 409–414.
Hulsen, M. 2000. *Language Loss and Language Processing: Three Generations of Dutch Migrants in New Zealand*. Nijmegen, The Netherlands: PhD dissertation, Nijmegen University.
Kootstra, G. J., Van Hell, J. G., & Dijkstra, T. (this volume). An interactive alignment perspective on code-switching in bilingual speakers.
Levelt, W. J. M. 1989. *Speaking: From intention to articulation*. Cambridge: MIT.
Myers-Scotton, C. 1993. Common and uncommon ground: Social and structural factors in codeswitching. *Language in Society*, 22: 475–503.