Bilingualism: Language and Cognition 12 (4), 2009, 447–462 © Cambridge University Press 2009 doi:10.1017/S1366728909990204

Triggered codeswitching between cognate languages*

MIRJAM BROERSMA
Radboud University Nijmegen

This study shows further evidence for triggered codeswitching. In natural speech from a Dutch–English bilingual, codeswitches occurred more often directly next to a cognate (or "trigger word") than elsewhere. This evidence from typologically related, cognate languages extends previous evidence for triggering between typologically unrelated languages. With their large proportion of trigger words, the data provide insight into which words can trigger codeswitches; proper nouns, cognate content words with good and moderate form overlap, and cognate function words all induced codeswitching. Further, this study extends the evidence for triggered codeswitching from speech with relatively little codeswitching to speech with a high codeswitching density. In contrast with earlier work, not only words directly following a trigger word but also words directly preceding one were codeswitched more often than other words, suggesting that the scope of triggered codeswitching depends on the frequency of trigger words and of codeswitches in the speech.

> "When I was still working, we mostly spoke English, because my son was home, my husband was home, and I was home. . . . And to avoid confusion, it was better to have one language. But now that it's just the two of us, now it's mixed."

Multilinguals have important choices to make whenever they open their mouth: which language to use in this setting, with this person, about this topic? As the above quotation from an elderly lady who had moved from The Netherlands to New Zealand 34 years earlier nicely illustrates, language choice can involve awareness of the appropriateness of a certain language or a certain type of language use, and conscious decisions. As the speaker reported, she used to aim for at least largely monolingual English speech when that seemed appropriate. However, in other situations she clearly made other choices, because in a conversation with a Dutch–English bilingual interviewer, this is what she actually said:

(1) *Zolang ik werkte praatten we meest Engels,* BECAUSE *mijn zoon was thuis, m'n man was thuis, en ik was thuis. . . . En om geen verwarring te hebben het was beter* TO HAVE ONE LANGUAGE. *Maar nu we met z'n beiden zijn, nou is het gemengd.*

"When I was still working, we mostly spoke English, because my son was home, my husband was home, and I was home. . . . And to avoid confusion, it was better to have one language. But now that it's just the two of us, now it's mixed."

* I am grateful to Kees de Bot for useful discussions over the years. Thanks to Madeleine Hulsen for making her Dutch–English data available, to Laurence Bruggeman, Sybrine Bultena, Lies Cuijpers, Marieke Pompe, and Matthias Sjerps for coding the data, and to three anonymous reviewers for helpful comments.

Language choices can come about in different ways (Kroll, Bobb and Wodniecka, 2006), some of them more under the speaker's conscious control than others. This paper addresses one type of codeswitching that is not under the speaker's control. It is argued that when a speaker is in a situation where she feels free to codeswitch, the actual codeswitches may not be consciously planned but may sometimes occur under the influence of cognates. Thus, while the fact that codeswitches occur might largely depend on social and pragmatic considerations (e.g., Blom and Gumperz, 1972; Myers-Scotton, 1993), the place where they occur can be influenced by the presence of cognates.

There is growing evidence that cognates are related to the occurrence of some codeswitches. The earliest observations on this relation were made by Michael Clyne (Clyne, 1967), who noticed that cognates and codeswitches seemed to co-occur relatively often in bilingual speech. He studied a large number of immigrant populations in Australia, and found many examples of co-occurring cognates and codeswitches in the speech of German–, Croatian–, Dutch–, Vietnamese–, Italian–, and Spanish–English bilinguals, and Hungarian–German–English and Dutch–German–English trilinguals (Clyne, 2003). He called this relation, where cognates facilitate codeswitching, TRIGGERING (Clyne, 1967, 1972, 1977, 1980, 2003).

The first statistical evidence for the relation between cognates (or TRIGGER WORDS) and codeswitches was provided by Broersma and De Bot (2006), with a study on Dutch – Moroccan Arabic bilinguals. That study showed that words that immediately followed on a cognate were significantly more often codeswitched than words that did not follow on a cognate. Also, words that were not adjacent to a cognate but that were part of the same basic clause

Address for correspondence:
Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, PO Box 9104, 6500 HE Nijmegen, The Netherlands
mirjam.broersma@mpi.nl

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

as a cognate were more likely to be codeswitched than words that were not in a basic clause with a cognate. Thus, cognates were shown to trigger codeswitching of the surrounding words.

Broersma and De Bot (2006) proposed that the triggering effect is the result of the selection of the cognate from the mental lexicon. Cognates, that is words similar in form and meaning in two languages, might be strongly connected in the mental lexicon. Processing cognates differs from processing non-cognates, as has been shown with bilinguals (Dijkstra, Grainger and Van Heuven, 1999; Schulpen, Dijkstra, Schriefers and Hasper, 2003; Lemhöfer and Dijkstra, 2004), as well as trilinguals (Van Hell and Dijkstra, 2002; Dijkstra and Van Hell, 2003), and their conceptual representations are more tightly connected than that of non-cognate translation pairs (De Groot and Nas, 1991; Van Hell and De Groot, 1998). The selection of a cognate might therefore lead to a change in the activation of both languages at the lexical level, such that the activation of the least active language gets boosted (cf., Grosjean, 1998; Paradis, 2004). If that language was strongly activated already, which is likely in a codeswitching setting, this extra activation can be enough to tip the balance, such that the next time a lemma is selected, it might be one from this language instead of the one that was spoken before. Thus, selection of a cognate increases the chance of codeswitching.

Triggered codeswitching was indeed found in a conversation among three young Dutch – Moroccan Arabic speaking men. Dutch and Moroccan Arabic are typologically unrelated languages. They belong to the Germanic and Semitic language family, respectively, and do not share many cognates; in the Broersma and De Bot (2006) corpus, 4.7% of the words (104 out of 2224) were cognates. It is possible that for languages that share fewer cognates, the impact of a cognate might be much greater than for languages that share many cognates. Therefore, in the present study, the relation between cognates and codeswitching is further explored with a lexically strongly related language pair: Dutch and English. Dutch and English are both West Germanic languages and share many cognates. In the present study, 71.4% of the words (2035 out of 2849) were cognates. Clyne (1977) found many examples of co-occurring cognates and codeswitches in the speech of Dutch–English bilinguals, and the present study aims to show with statistical evidence whether cognates trigger codeswitching in this cognate language pair like they did for the typologically dissimilar language pair.

In the Dutch – Moroccan Arabic corpus, all cognates were nouns, and most of them proper nouns. Although proper nouns presumably need to be stored and processed as any other lexical item (Cutler, McQueen and Robinson, 1990), they form a quite specific subset and there is some evidence that they might sometimes be treated differently than other words, for example by aphasic patients (Van Lancker and Klein, 1990). Therefore, it would be useful to investigate the triggering effect of cognates that are not proper nouns. Further, it is possible that nouns and function words might vary in their triggering potential. The amount of form overlap of cognates might also play a role; e.g., it is possible that only cognates with nearly perfect form overlap can trigger a codeswitch. As Dutch and English share many cognates, those languages are very suitable for investigating the triggering effect of different types of cognates.

Clyne (1967, 1972, 1977, 1980) proposed that the words that are most likely to be codeswitched are the words immediately preceding and following a cognate. Broersma and De Bot (2006) found that in the Dutch – Moroccan Arabic corpus, words immediately following a cognate had an increased chance of being codeswitched indeed, but this was not the case for words immediately preceding a cognate. They also extended the prediction to words that are further away from the cognate; they showed that words within the same basic clause as the cognate, even if they were not immediately bordering on the cognate, still had an increased chance of being codeswitched. The present paper aims to further investigate the scope of triggered codeswitching: does triggered codeswitching only concern words that follow a cognate, or does it affect words that precede a cognate too, and what is the role of the clause level for triggered codeswitching? The answers to these questions might depend on the frequency of codeswitches in the data: in speech with a dense codeswitching pattern, codeswitches may often occur close to the cognate, which might diminish the value of a wider view. The Dutch–English data contain many more codeswitches than the Dutch – Moroccan Arabic corpus: in the Dutch–English data, 17.0% of the words were codeswitched (138 out of 814 non-cognate words), and 34.6% of the basic clauses contained a codeswitch (106 out of 306 basic clauses); in the Dutch – Moroccan Arabic data, 2.8% of the words were codeswitched (60 out of 2120 non-cognate words), and 16.4% of the basic clauses contained a codeswitch (52 out of 318 basic clauses). Thus, the present paper also aims to investigate the scope of triggering in a dense codeswitching situation.

Finally, the study aims to extend the evidence for triggered codeswitching to another population and another conversational setting. Whereas the informants in Broersma and De Bot (2006) were young men who had been born in The Netherlands or had moved there during childhood, the informant in the present study was an elderly lady who had immigrated to New Zealand as an adult, and while the Dutch – Moroccan Arabic corpus consisted of self-recorded conversation among three friends, the present data consist of a conversation between the informant and a researcher.

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

1. Materials

1.1 The informant

The data consist of an interview with a female speaker, who moved from The Netherlands to New Zealand in 1961 with her husband and son, when she was 39 years old. Since that time she had visited The Netherlands four times for several months on each occasion. The interview took place in 1995, when she was 73 years old and had been living in New Zealand for 34 years.

The interview was carried out as part of a study into language loss among three generations of Dutch migrants in New Zealand (Hulsen, 2000). It took place in the informant's home, and lasted 24 minutes. The interview was conducted by a Dutch female, who asked the informant about her experiences around her immigration, life in New Zealand, visits to The Netherlands, language use, and attitude towards the use of Dutch and English. The tone of the conversation was informal, and the informant often deviated from the topics that the interviewer introduced. Although the informant was aware that the interviewer knew English well, the interviewer spoke only Dutch during the interview. She limited herself to questions and short responses, intended to elicit spontaneous, running speech from the informant. The informant was not instructed about language choice and not aware that her language use would be evaluated. Only the informant's speech is examined here.

The informant indicated that she had learned English in The Netherlands and did not experience much difficulty with the language when she first arrived in New Zealand, although she still occasionally encountered words that she did not know. Soon after the immigration, English became the main language within the home. She also spoke English with neighbors, friends, and at work. Recently, her use of Dutch had increased again. She suggested that this might be because with older age, her memories about The Netherlands became more important to her. Since their son had left home, she and her husband spoke both Dutch and English with each other. She also used both languages with her son and with her grandchildren. The grandchildren, however, knew little Dutch. She was a member of the local Dutch association, but she expressed mixed feelings about it and did not often join their meetings.

During the conversation, the informant very regularly codeswitched between Dutch and English. Note that codeswitching is typically found in informal conversations within the peer group. An interview with a researcher who uses one language only is not a setting in which codeswitching is typically expected. First-generation immigrants who have lived outside their native language environment for a long time, however, have been found to codeswitch in such settings, especially when they aim to speak in their native language (e.g., Clyne, 1967, 1977).

1.2 Characterizing the data: codeswitching and transference

Before turning to the analysis of triggered codeswitching in the data, a description of the types of codeswitching and transference in the data is provided. The base language in the largest part of the conversation was Dutch and contained only few fully English sentences. The informant's speech contains many cognates and many codeswitches: 71.4% of the words (2035 out of 2849) were cognates, and 17.0% of all the non-cognate words (138 out of 814) were codeswitched. At the clause level, 57.8% of the basic clauses (177 out of 306) contained a cognate, and 34.6% of the basic clauses (106 out of 306) contained a codeswitch.

This count of the number of codeswitches, however, does not do justice to the complexity of the informant's speech. Her speech is characterized by the combination of elements from the two languages in varying ways, at the syntactic, lexical, morphological, and phonological levels, leading to so called COMPROMISE FORMS (Hasselmo, 1961; Clyne, 1967, 1977). For entire passages of speech, it is therefore impossible to decide which language she is speaking, or which language any element of her speech comes from. Such MARGINAL PASSAGES have been reported to occur in the speech of first-generation elderly immigrants (Hasselmo, 1961; Clyne, 1967, 1977; De Bot and Clyne, 1989). This type of language use can arise more easily in languages that are similar on many levels, like Dutch and English, and the overlap on different non-lexical levels might facilitate codeswitching (Clyne, 2003; Broersma, Isurin, Bultena and De Bot, 2009; De Bot, Broersma and Isurin, 2009). In order to grasp the complexity of the informant's language use, it is necessary to know something about the types of codeswitching as well as the types of transference in her speech. Before moving on to the analysis of triggered codeswitching, the patterns in the data are described using two frameworks that are very useful for this goal: first, Muysken's (2000) classification of codeswitches as insertion, alternation, and congruent lexicalization, and second, Clyne's (2003) classification of syntactic, semantic, morphological, and phonological transference.

To illustrate these phenomena, examples from the data are provided. Note, however, that these examples cannot be fully discussed until after the operationalization of cognates is presented in section 2.1. For now, the examples are presented as if it could be determined for each word which language it came from. For cognates, however, that can never be unambiguously determined. After the operationalization of cognates is discussed, the examples will therefore be re-interpreted, taking this ambiguity into

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

account. Due to the abundant presence of cognates in the data, the examples will be evaluated rather differently then. For now, each word is assigned to one language. The language of each item is determined on the basis of the phonological form and, where necessary, the phonetic realization. Dutch words are given in italics and English words in small capitals.

Insertion, alternation, and congruent lexicalization

Muysken (2000) classifies codeswitches as insertions, alternations, and congruent lexicalizations. All three types of codeswitches are present in the data. First, there are many insertional codeswitches in the data in the form of single words from one language occurring in a sentence frame from the other language. In example (2), an English word is embedded in a Dutch context, and the reverse is the case in (3).

(2) *De enige ding wat we hadden was dat linker en rechter verkeer. Dat was onze probleem.* BECAUSE *als we dachten naar huis te gaan dan gingen we net die andere kant op. We zaten aan de verkeerde kant van de bus te wachten.*

"The only thing that we had was that left and right traffic. That was our problem. Because if we thought we were going home, then we went just the other way. We were waiting on the wrong side of the bus."

(3) SO IF HE HAD A PROBLEM, HE WAS JUST GOING TO SEE HER AND EXPLAIN HIM AND HE WAS *klaar*.

"So if he had a problem, he was just going to see her and (she'd) explain (it to) him and he was ready."

In other cases of insertional codeswitching, a larger constituent is codeswitched. In example (4) an English Prepositional Phrase and in (5) a Dutch Noun Phrase are embedded in a sentence frame from the other language.

(4) *Je concentreert meer* ON WHAT WAS THAN WHAT IS COMING.

"You concentrate more on what was than what is coming."

(5) I DON'T THINK THEY HAVE A *goeie reglement*.

"I don't think they have a good regulation."

Second, in (6), a relative clause is codeswitched, which is an example of an alternational codeswitch.

(6) *Want ik had nog land daar* THAT I WANTED TO SELL.

"Because I still had land there that I wanted to sell."

There are many examples where a codeswitch occurs between sentences, as in (7) and (8), which are also alternations.

(7) CAN *je je voeten op zetten*. WOULD BE QUITE NICE AT THE MOMENT.

"You can put your feet on (it). Would be quite nice at the moment."

(8) *Ik zei:* ELLEN NEED A BIGGER DUVET.

"I said: Ellen need(s) a bigger duvet."

Finally, other cases of codeswitching are less transparent. In (9), codeswitching takes place within several levels of the sentence structure, and it seems that both languages contribute to the grammatical structure of the sentence. This is an example of congruent lexicalization.

(9) *Maar het was juist* TO HAVE *niet de vrees* BEHIND YOU *dat de Russen wouden komen*.

"But it was just not to have the fear behind you that the Russians would come."

Transference

The informant's speech is characterized by transference, that is, by the influence of one language on the other (Odlin, 1989; Clyne, 2003), on all levels of speech. Below, examples of syntactic, semantic, morphological, and phonological transference in the data are discussed, following Clyne's (2003) classification. (Note that there are other forms of transference, e.g., at the discourse level (Odlin, 1989), that are not discussed here.) There is more evidence for transference from English to Dutch in the data than vice versa. There are no clear examples of syntactic and semantic transference from Dutch to English. This may be because the base language of this conversation was Dutch, with only few sentences in English.

Syntactic transference

There is experimental evidence that syntactic structures are shared between languages (Hartsuiker, Pickering and Veltkamp, 2004; Desmet and Declercq, 2006), and indeed bilingual speakers often show transference of syntactic structures of one language to the other (Clyne, 2003). In the present study, the informant regularly uses English constructions in combination with Dutch words. Dutch and English word order overlap to some extent, but there are many differences, as demonstrated in the following examples.

(10) *Later ik naaide voor mensen*.

"Later I sewed for people."

The basic word order in Dutch is SVO. However, if a sentence adverb is added in sentence-initial position, the order of subject and verb is inverted. In example (10), the adverb "later" precedes the subject ("ik"), which nevertheless precedes the verb "naaide". This word order

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

is correct in English but not in Dutch. The correct order would have been "Later naaide ik voor mensen".

(11) *Je kunt hebben dan een maaltijd*.

"You can have a meal then."

In Dutch, the infinitive should appear at the end of the sentence. In (11), the infinitive "hebben" is placed directly after the finite verb "kunt" and before the object "een maaltijd", following the English word order. The correct order would have been "Je kunt dan een maaltijd hebben".

Semantic transference

When learning a second language, meanings from native language words are often transferred to second language words and vice versa (Dong, Gui and MacWhinney, 2005). The data show many instances of semantic transference, where the meaning of an English word is transferred to a Dutch word. The words involved share some aspects of their meaning, but not others. In example (11) above, the verb "hebben" ("to have") is used with the meaning "to get". Whereas the English verb can carry this meaning, the Dutch verb cannot. The appropriate verb here would have been "krijgen".

(12) *Begin maart zijn we naar een Friese dag geweest juist uit Pyronia*.

"Beginning of March we have been to a Frisian day just outside Pyronia."

In (12) "juist uit" is not a correct Dutch expression. Although the separate words carry the same meaning as the English words "just" and "outside", the combination of "juist" and "uit" suggests a temporal meaning for the component "juist" and an outward movement for the component "uit". An appropriate expression would have been "net buiten".

(13) *Wat was even meer gezeur*.

"Which was even more hassle."

In (13), Dutch "even" is used with the meaning of "even more". The Dutch and English words "even" share the meaning of "divisible by two". Dutch "even" has many other meanings, but it cannot be used to express "even more". An appropriate expression would have been "nog meer".

(14) *Dat was een andere twee maanden*.

"That was another two months."

The Dutch word "andere" means "other". The sentence in (14) is a direct translation of the English construction "another two months", which carries the meaning of "more" ("two more months"). The Dutch word "andere" does not have this meaning. Apart from the semantic transference, the sentence also follows the English construction and deviates from the Dutch structure used to express "other": "Dat waren twee andere maanden." ("Those were two other months."). A correct way to express the meaning of "two more months" would have been: "Dat waren nog twee maanden." ("That was two more months.").

Morphological transference

There are some cases of mixed morphology, where lexemes consist of both Dutch and English morphemes. In all these cases, the bound morphemes come from Dutch, and free morphemes come from either language. In terms of the Matrix Language Frame model (Myers-Scotton, 1997, 2006), Dutch would be called the Matrix Language, and English the Embedded Language.

(15) *Dat was landings*MONEY.

"That was landing money."

In (15), the word "landingsmoney" (i.e., the money that governments required immigrants to possess when they entered the country) contains both Dutch and English morphemes. The part "landing" is almost homophonous in the two languages, only differing in the first vowel (Dutch: /ɑ/, English: /æ/), and is here pronounced with the Dutch vowel. The part "money" is exclusively English. The Dutch linking morpheme "-s-" is used to connect the two parts of the compound.

(16) *Ik bleef maar door*TRAVEL*en*.

"I just kept travelling on."

In (16), "doortravelen" is a combination of the English stem "travel" and two Dutch morphemes. The English stem is preceded by the Dutch morpheme "door", indicating continuity, and followed by the Dutch infinitive marker "-en".

Phonological transference

Bilingual speakers often pronounce phonemes differently than monolingual speakers of that language do (Flege, Schirru and MacKay, 2003), either because their phonological representations differ, or because phonological representations of one language are activated while speaking the other language (Roelofs and Verhoeff, 2006). The phonology of the native language can influence that of the second language, but the reverse is found as well (Flege and Eefting, 1987; Bullock and Toribio, 2009). Being in a codeswitching setting also influences pronunciation, in widely varying ways, from divergence to convergence, and from interference to hypercorrection (Bullock and Toribio, 2009). In the present study, the informant also shows Dutch influences in her English phonology. Although her English pronunciation seems generally good, she sometimes pronounces English words, and especially particular phonemes, in an atypical way which is consistent with

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

Dutch phonology. Some phonemes that do not occur in Dutch are pronounced as a phonetically close Dutch phoneme.

For example, where English distinguishes two open mid-front unrounded vowels /æ/ and /ɛ/, Dutch has only an /ɛ/ in this part of the vowel space, which makes the contrast difficult to perceive for Dutch–English bilinguals (Broersma, 2005). The informant regularly pronounces the /æ/ in English words as /ɛ/, as in the following example:

"apparently" (Dutch: "blijkbaar")	here:	/əˈpɛrəntlɪ/
	English:	/əˈpærəntlɪ/

Dutch has no dental fricatives, and the informant often pronounces these as stops, replacing /θ/ with /t/ and /ð/ with /d/:

"they" (Dutch: "zij")	here:	/deɪ/
	English:	/ðeɪ/

When Dutch and English word forms overlap, in many cases, it cannot be determined whether a certain phoneme replacement is made because of difficulty with this particular phoneme, or analogous to the Dutch translation equivalent. In the following example, /ð/ might have been replaced with /d/ analogous to the Dutch translation:

"that"	here:	/dɛt/
	English:	/ðæt/
	Dutch:	/d*a*t/

Similarly, in the next example, the stress pattern and the second and third vowel are in line with the English word. The fricative /θ/ is replaced with /t/, which is analogous to the Dutch translation, but also in line with the general pattern described above. The first vowel is pronounced as /ɛ/. Note that the speaker does not replace the English /æ/ with the vowel from the Dutch equivalent, /ɑ/, but with /ɛ/ instead, which is phonetically closer. There is no suffix "-e", which would be required in the given context in Dutch.

"Catholic"	here:	/ˈkɛtəlɪk/
	English:	/ˈkæθəlɪk/
	Dutch:	/k*a*toːˈliːk/

The following example is very similar to the previous one. It is also pronounced with the English stress pattern. (Note that there is a similar Dutch word with the same stress pattern, but this is an adjective, meaning "mathematical".) The fricative /θ/ is replaced with /t/, which is again analogous to the Dutch translation, but also in line with the general pattern. The two occurrences of /æ/ are not replaced with the vowels from the Dutch equivalent /ɑ/, but with the phonetically closer /ɛ/.

"mathematics"	here:	/mɛtəˈmɛtɪks/
	English:	/mæθəˈmætɪks/
	Dutch:	/m*a*təːmaːˈtiːk/
	Dutch "mathematical":	/m*a*təːˈmaːtiːs/

Often, it cannot be determined how morphological and phonological transference contributed to the realization of a word. In the following example, the stress pattern and the first three vowels are in line with the English pronunciation. However, the fourth vowel is in line with the Dutch pronunciation. (Note that unstressed vowels are not necessarily reduced in Dutch.) The word ends with the Dutch suffixes "-s-e". Thus, the word might consist of only Dutch morphemes, while the stress pattern and the first three vowels are an influence from English, or it may consist of a combination of an English part (/əˈmɛrɪkan/), with some characteristics of the phonological realization of the Dutch word, and the Dutch suffixes "-s-e".

"American"	here:	/əˈmɛrɪkansə/
	English:	/əˈmɛrɪkən/
	Dutch:	/aːmeːriːˈkaːnsə/

2. Method

The above description shows that the present materials form a rich set of bilingual speech data, containing a complex mix of English and Dutch. With their dense and complex pattern of codeswitching, and their large number of cognates, they provide an interesting test for the triggering theory. Three main questions are addressed. First, do cognates trigger codeswitches for the typologically related and strongly cognate languages Dutch and English like they did for the typologically unrelated languages Dutch and Moroccan Arabic? Second, what types of cognates can trigger codeswitches? Third, how far does the triggering influence of cognates reach and which words are affected by the occurrence of cognates?

Cognates play a central role in the triggering theory. In order to test the triggering theory, it needs to be defined which words are considered to be cognates. The operationalization of this term requires some choices and therefore, to avoid confusion, the term "trigger words" is used instead of "cognates". After defining trigger words, the examples of codeswitches that were given in section 1.2 above are reconsidered. Labeling words as trigger words changes the way these examples are interpreted. Finally, the procedure section outlines how the data were analyzed for the occurrence of triggered codeswitching.

2.1 Trigger words

Trigger words are words that overlap both in form and in meaning in two languages. Following the definition

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

proposed by Broersma and De Bot (2006), trigger words comprise bilingual homophones and proper nouns, allowing for small differences in phonological form.

In the definition by Clyne (2003), apart from bilingual homophones and proper nouns, the category of trigger words also contains lexical transfers. Lexical transfers are items that belong to one language, but have become part of the lexicon of the other language for the individual speaker. What is an individual speaker's codeswitch on its first occurrence might become entrenched in the language of the wider community after repeated use (Backus, 1996, 2009; Winford, 2009). Lexical transfers are somewhere between these endpoints: they are part of the language for the individual speaker, but not for the wider community (yet). The individual history of bilingual speakers, however, which determines the distinction between codeswitches and borrowings for that speaker may vary widely (Odlin, 2009). As argued by Broersma and De Bot (2006), there is no clear-cut way to determine which words have been incorporated into another language for an individual speaker. Acknowledging that lexical transfers may play a role in triggered codeswitching, the category of lexical transfers is therefore abandoned here in order to operationalize trigger words.

English and Dutch have many translation equivalents that are fully homophonous, in the sense that their IPA notations for the two languages are identical. Such trigger words include, for example:

/ɪn/ – /ɪn/ "in" – "in"
/ən/ – /ən/ "an" – "een"

Other translation pairs only differ in phonemes that are phonetically so close that it is impossible for the listener to determine with any certainty what the intended phoneme is. These words are also considered to be trigger words. Some examples from the data are:

/aʊt/ – /œyt/ "out" – "uit"
/maɪ/ – /mɛi/ "my" – "mij"
/miː/ – /mɛi/ "me" – "mij"

Some words become (nearly) homophonous as they are often pronounced in a reduced form in regular speech. Such words are also considered trigger words. Examples include:

/t/ (/ɪt/) – /t/ (/ət/) "it" – "het"
/jə/ (/juː/) – /jə/ "you" – "je"
/dɪs/ (/ðɪs/) – /deːs/ (/deːzə/) "this" – "deze"

Some words are not homophonous in citation form, but become very similar in regular speech and in the informant's pronunciation. For example, word-final stops are often unreleased (Byrd, 1993). Combined with the informant's phonological transference of /æ/ being pronounced as /ɛ/, this makes the words in the following examples quite similar:

/ɛnd/ (/ænd/) – /ɛn/ "en" – "en"
/hɛv/ (/hæv/) – /hɛp/ "have" – "heb"

For the present study, six judges decided independently from one another for each word in the data whether they considered it as a trigger word or not, following the definition given above, allowing for some variation in phonological form, and taking into account that some translation pairs are more similar in this speaker's pronunciation and in running speech than in citation form. Results of the judges' consistency are reported below. The proportion of trigger words in the data is very large. Out of a total of 2849 words, 2035 words (71.4%) are trigger words. Note that among the trigger words are highly frequent function words, such as "in", "an", "it", "and", etc.

Due to the large number of trigger words, long stretches of speech consist of trigger words only. Even though the morphological, phonological and phonetic realization may suggest that the speech comes from one particular language, this does not necessarily imply that the lemmas that were selected from the mental lexicon to produce the utterance were indeed part of that language. The essence of the notion of trigger words is that it cannot be unambiguously determined which language they come from. To represent the data as clearly as possible, in the transcribed examples, for most trigger words it is indicated if the realization is more in line with the Dutch (italics) or with the English form (small capitals). The spelling follows that language too. This is not meant to suggest that the trigger word belongs to that language. (In some cases, the pronunciation of the trigger word matches both languages equally well, in which case no language is indicated.)

2.2 Codeswitching in the data reconsidered

The examples of codeswitching given above are reconsidered here, now also taking into account which words are trigger words. By definition, the language of a trigger word cannot be determined, and trigger words are never counted as a codeswitch. Of course this is an underestimation of the actual number of codeswitches, but in order to determine the effect of trigger words on codeswitches, a strict separation of these two categories is necessary.

As the majority of the words in the data are trigger words, for convenience, trigger words are NOT underlined, and non-trigger words are underlined in all the following examples. Note that the Dutch and English spelling of trigger words may differ considerably, without indicating poor overlap in phonological form. Dutch words are given in italics and English words in small capitals. For trigger

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

words, italics and small capitals indicate which language the pronunciation matches best.

With the trigger words taken into account, the patterns described in section 1.2 are not so clear anymore. For example, it can no longer be claimed that the entire relative clause in (6) is codeswitched (cf. (6′)). Taking trigger words into account, examples (7) and (8) do not even contain a codeswitch (cf. (7′) and (8′)), and example (4) contains only trigger words and no single item that can be ascribed to one particular language (cf. (4′)). Thus, the categorization of words as trigger words (the language of which cannot be determined and which cannot be codeswitched) versus other words is rigorous and necessarily means that some information is lost. Arguably, it leads to an underestimation of the amount of codeswitching. This categorization is, however, crucial for the current study, as it makes the statistical testing of the triggering theory possible.

(2′) *De enige ding wat we hadden was dat linker en rechter verkeer. Dat was onze probleem.* BECAUSE *als we dachten naar huis te gaan dan gingen we net die andere kant op. We zaten aan de verkeerde kant van de bus te wachten.*

"The only thing that we had was that left and right traffic. That was our problem. Because if we thought we were going home, then we went just the other way. We were waiting on the wrong side of the bus."

(3′) SO IF HE HAD A PROBLEM, HE WAS JUST GOING TO SEE HER AND EXPLAIN HIM AND HE WAS *klaar*.

"So if he had a problem, he was just going to see her and (she'd) explain (it to) him and he was ready."

(4′) *Je concentreert meer* ON WHAT WAS THAN WHAT IS COMING.

"You concentrate more on what was than what is coming."

(5′) I DON'T THINK THEY HAVE A *goeie reglement*.

"I don't think they have a good regulation."

(6′) *Want ik had nog land daar* THAT I WANTED TO SELL.

"Because I still had land there that I wanted to sell."

(7′) CAN *je je voeten op zetten*. WOULD BE QUITE NICE AT THE MOMENT.

"You can put your feet on (it). Would be quite nice at the moment."

(8′) *Ik zei:* ELLEN NEED A BIGGER DUVET.

"I said: Ellen need(s) a bigger duvet."

(9′) Maar het was juist TO HAVE niet de vrees BEHIND [YOU/je] dat de Russen wouden komen.

"But it was just not to have the fear behind you that the Russians would come."

2.3 Analysis

In order to test the triggering theory, the interview was transcribed and coded for the occurrence of trigger words and codeswitches, and statistical analyses were performed to assess the relation between the two.

Levels of analysis

To investigate the scope of triggered codeswitching, triggering was assessed both at the word level and at the clause level. The original triggering theory, as it was presented by Clyne (1967, 1972, 1977, 1980), considers codeswitching at the word level, and in a linear way. In this approach, a word is considered to be codeswitched when it differs in language from the previous word, regardless of their grammatical relation, and regardless of when a switch back occurs. Thus, single lexical items can also constitute a codeswitch. Clyne proposes that trigger words could facilitate codeswitching of directly preceding and of directly following words.

The adjusted version of the triggering theory presented by Broersma and De Bot (2006) considers codeswitching at the clausal level. Broersma and De Bot (2006) propose that triggering may not be limited to words directly bordering on a trigger word, but may extend to words that are part of the same basic clause. The level of analysis here is the basic clause. A basic clause is considered to contain a codeswitch when it contains words from two languages, or when it contains words from a language different than that in the previous basic clause. Note that no language (e.g., base or matrix language) is assigned to the basic clause as a whole; rather, it is assessed whether a clause contains non-trigger words from either language.

In the Dutch – Moroccan Arabic data (Broersma and De Bot, 2006), words that preceded a trigger word did not have an increased chance of being codeswitched. Words that followed a trigger word, however, did: they were codeswitched more often than words that did not border on a trigger word. In addition to that, words that did not directly border on a trigger word but that were in the same basic clause as a trigger word also had an increased chance of being codeswitched.

To make the prediction about triggering at the clause level more specific, in the current study it was also tested how often a codeswitch concerned the entire basic clause containing the trigger word (i.e., all the non-trigger words in that clause). In the Dutch – Moroccan Arabic data, 39.6% of the codeswitches occurred between basic clauses, and 60.4% within basic clauses. Of all the codeswitches that the triggering theory explained at the clause level, only 52.4% were intra-clausal codeswitches. The other 47.6% were inter-clausal codeswitches. In those

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

cases, the codeswitch occurred at the beginning of the basic clause that contained the trigger word, as in (17), where the first basic clause is in Dutch and the second in Moroccan Arabic with a trigger word in the second basic clause.

(17) *Waarom niet?* L-*familie* dyal-ek, tšuf-ha.

"Why not? You see your family."

Thus, when a trigger word triggered a codeswitch, in almost half the cases this resulted in a switch of an entire basic clause rather than only part of it. This explains why the clause-level analysis had a substantial contribution to make on top of the predictions of the word-level analysis.

In the present study, triggering is assessed both at the word level (for words preceding and following a trigger word) and at the clause level (for intra- and inter-clausal codeswitches).

Procedure

In order to test triggering at the word level and at the clause level, different testing procedures were required. Differences concern both the measuring unit and the definition of what constitutes a codeswitch. The procedure was as described by Broersma and De Bot (2006). First, for both tests, each word was categorized either as a trigger word or a non-trigger word (following the operationalization described above), and the language of each non-trigger word was determined.

For the word-level analysis, the data were divided into conversational turns, lasting as long as the informant was speaking without interruption from the interviewer. The first non-trigger word of each conversational turn served as a reference point to determine the starting language of that turn. For each non-trigger word it was determined whether it directly preceded or followed a trigger word, and whether it was codeswitched, and the number of words in each category was counted. A word was considered to be codeswitched when it was part of a different language than the previous non-trigger word.

For the clause-level analysis, on the other hand, the data were divided into basic clauses, defined as utterances that contain maximally, but not minimally, one main verb (Levelt, 1989). For every basic clause it was determined whether it contained one or more trigger words, or a codeswitch, and the number of basic clauses in each category was counted. A basic clause was considered to contain a codeswitch when it contained words from two languages (an intra-clausal codeswitch), or when it contained words from a language different than that in the previous basic clause (an inter-clausal codeswitch). If the previous basic clause contained only trigger words, the language was compared to that of the nearest preceding basic clause with a non-trigger word. If a basic clause without an intra-clausal codeswitch followed a basic clause with an intra-clausal codeswitch, it could not be determined whether there was a codeswitch relative to this preceding basic clause; therefore, in such cases the former clause was left out of the analysis. Basic clauses containing only trigger words could not contain a codeswitch, and were left out of the analysis. Note that whereas for the word level, analyses proceeded in a left-to-right fashion, word order was not crucial for the clause level. Note also that no language (e.g., base language or matrix language) was assigned to the entire basic clause; rather, the presence of non-trigger words from either language was assessed.

For both levels of analysis, two statistical tests were used to investigate the relationship between trigger words and codeswitches. Both tests assess in similar ways whether two variables are independent of each other. The first is the χ^2 test for independence, which is the test most commonly used for this purpose. However, it may be less reliable if there is a small value in one of the cells, or if the marginal is very uneven. Therefore, Fisher's Exact test, which is more accurate in such cases, is reported as well. Fisher's Exact test yields a one-sided probability, here indicated as "P".

Inter-rater reliability

Each word was categorized as either a trigger word or a non-trigger word. As the examples in section 2.1 illustrate, this is not always a straightforward decision, and some subjectivity cannot be avoided. The analyses were based on the coding of the first judge (the author of this paper). In order to assess the reliability of this judge's coding, five other judges coded the entire data set independently from each other as a control. All judges were native speakers of Dutch who were proficient in English as a second language. The first judge transcribed the audio recording of the interview. All judges listened to the audio recording and added their codes to the transcription.

It was first determined how strong the agreement among all six judges was. The inter-rater reliability (calculated with a Two-Way Random Effects Model, Type Absolute Agreement, Average Measures Interclass Correlation Coefficient (ICC)) was high ($ICC = .914$, $p < .001$). Next, the correlation between the first judge's coding and each of the five other judges' coding was determined. This correlation was also high (averaged Pearson $r\ (2552) = .6248$, $p < .01$). These results warrant the further analyses based on the first judge's classification of trigger words and non-trigger words.

3. Results

3.1 Word level

First, triggering was evaluated at the word level. The word-level analysis is illustrated below with some examples.

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

(Superscript numbers can be ignored for now and will be clarified later.) In (18), trigger words and codeswitches follow each other rapidly and in a regular pattern: all the words that are not trigger words are codeswitched, and after each trigger word follows a codeswitch. The sentence starts with a trigger word. As the sentence follows a fully Dutch fragment, the word "apparently" is a codeswitch from Dutch to English. Then follows the trigger word "deze", a codeswitch back to Dutch, the trigger word "was", a codeswitch to English, a proper noun ("R.", the name of the informant), a codeswitch to Dutch, a codeswitch to English, and another trigger word.

(18) [*En*/AND] APPARENTLY[1] *deze mevrouw*[2] was CALLED[3] R., *net*[4] LIKE[5] ME.

"And apparently this lady was called R., just like me."

In (19), the proportion of trigger words is very high. Among the words that can be unambiguously attributed to one language, most are Dutch, and so is the phonology of the largest part of the fragment. In the first sentence, a single English word, "because", occurs. In the second sentence, the constituent "to have one language" seems to come from English, but only the word "language" is a non-trigger word. A word was considered to be codeswitched 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. By definition, however, it is always in the same speech turn. For example, in (19) "because" is codeswitched relative to "praatten", "thuis" relative to "because", and "language" relative to "verwarring".

(19) *Zolang ik werkte praatten*[6] *we meest Engels*, BECAUSE[7] *mijn zoon was thuis*[8]*, mijn man was thuis*[9] *en ik was thuis*[10] *en mijn nichtje*[11] *was hier toen. . . . En om*[12] *geen*[13] *verwarring*[14] *te hebben, het was beter* TO HAVE ONE LANGUAGE[15].

"As long as I worked, we spoke mostly English, because my son was home, my husband was home, and I was home, and my niece was here then. . . . And to have no confusion, it was better to have one language."

Example (20) contains only trigger words and Dutch words. As it also follows a Dutch sentence, there is no codeswitching in this fragment.

(20) *En ik moest naar*[16] *Roermond, want*[17] *ik kon niet in Tegelen blijven*[18]*. Daar is niets*[19] *waar je kunt blijven*[20].

"And I had to go to Roermond, because I couldn't stay in Tegelen. There is nothing where you can stay."

Table 1. *Number of words that are codeswitched, number of words that are not codeswitched, and percentage of words that are codeswitched; split by words that follow a trigger word and words that do not border on a trigger word.*

Following a trigger word	Codeswitch		
	Yes	No	% Yes
Yes	31	68	31.3
No	3	56	5.1

At the word level, the triggering hypothesis as it was originally presented (Clyne, 1967, 1972, 1977, 1980) would predict that words directly following or directly preceding a trigger word have a greater chance of being codeswitched than words that are not adjacent to a trigger word and, further, that words between two trigger words have a greater chance of being codeswitched than words that are adjacent to a trigger word on only one side.

To test the first prediction, all words directly following (and not preceding) a trigger word were compared to all words that were not adjacent to a trigger word. The words "net"[4] (in example (18)) and "language"[15] (in example (19)) are cases of words following a trigger word that are codeswitched, while "om"[12] (in (19)) is an example of one that is not codeswitched. The word "geen"[13] is an example of a word that is not adjacent to a trigger word, and that is not codeswitched. Table 1 shows that the percentage of codeswitches is much higher for words that follow a trigger word than for words that are not adjacent to a trigger word, with 31.3% versus 5.1%. This difference is statistically significant ($\chi^2 = 15.06$, $p < .0001$, $P < .0001$). Thus, words that directly follow a trigger word have an increased chance of being codeswitched.

To test the second prediction, all words directly preceding (but not following) a trigger word were compared to all words that were not adjacent to a trigger word. The word "like"[5] is an example of a word preceding a trigger word that is codeswitched, "verwarring"[14] is an example of one that is not codeswitched. Table 2 shows that the percentage of codeswitches is higher for the words that precede a trigger word than for the words that are not adjacent to a trigger word, with 22.5% versus 5.1%. This difference is statistically significant ($\chi^2 = 8.42$, $p < .01$, $P < .01$). Thus, words that directly precede a trigger word have an increased chance of being codeswitched too.

Both words that directly follow and words that directly precede a trigger word have an increased chance of being codeswitched, with, respectively, 31.3% and 22.5% being codeswitched. Comparing these two categories, there is no statistical difference in the proportion of codeswitches ($\chi^2 = 1.96$, $p > .1$, $P > .1$). Thus, words following a

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

Table 2. *Number of words that are codeswitched, number of words that are not codeswitched, and percentage of words that are codeswitched; split by words that precede a trigger word and words that do not border on a trigger word.*

Preceding a trigger word	Codeswitch		
	Yes	No	% Yes
Yes	23	79	22.5
No	3	56	5.1

Table 3. *Number of words that are codeswitched, number of words that are not codeswitched, and percentage of words that are codeswitched; split by words that border on a trigger word on two sides and words that border on a trigger word on one side.*

Bordering on a trigger word on	Codeswitch		
	Yes	No	% Yes
two sides	81	227	26.3
one side	54	147	26.9

trigger word and words preceding a trigger word have the same chance of being codeswitched.

To test the third prediction, all words between two trigger words were compared to all words that either only preceded or only followed a trigger word. The words "apparently"[1], "mevrouw"[2], "called"[3], "because"[7], and the first occurrence of "thuis"[8] are examples of words between two trigger words that are codeswitched. The second and third occurrences of "thuis"[9,10] and "nichtje"[11] are examples of words between two trigger words that are not codeswitched, and so are all the non-trigger words[16–20] in example (20). Table 3 shows that the percentage of codeswitches is virtually identical for words that occur between two trigger words and words that border on a trigger word on one side only. There is no statistical difference between the two categories ($\chi^2 = 0.02$, $p > .8$, $P > .5$). Thus, although bordering on a trigger word on one side increases the chance of being codeswitched, this chance is not further increased by bordering on a trigger word on the other side as well.

3.2 Clause level

Next, triggering was evaluated at the clause level. The data were divided into basic clauses and for each basic clause it was determined whether it contained a trigger word and whether it contained a codeswitch. This is illustrated with the same examples as those discussed above. Example (18) consists of one basic clause which contains trigger words as well as a codeswitch (as it contains Dutch and English words).

Example (19) consists of nine basic clauses listed in (19′). All these basic clauses contain several trigger words. Basic clauses [1] and [8] contain only trigger words and are therefore left out of the analysis. Basic clause [3] contains a codeswitch (as it contains both an English and a Dutch word), and so does [9] (as it contains an English word whereas the previous basic clause with non-trigger words, [7], contains Dutch words). Basic clause [4] is left out of the analysis, because as the previous basic clause contains two languages, it cannot be determined whether the Dutch word in [4] should be considered as a codeswitch or not. Basic clauses [2,5,6,7] contain no codeswitches.

(19′) [1] *zolang ik werkte*

[2] *praatten we meest Engels*

[3] BECAUSE *mijn zoon was thuis*

[4] *mijn man was thuis*

[5] *en ik was thuis*

[6] *en mijn nichtje was hier toen*

[. . .]

[7] *en om geen verwarring te hebben*

[8] *het was beter*

[9] TO HAVE ONE LANGUAGE

Example (20) consists of six basic clauses, as shown in (20′). Basic clauses [1]–[4] contain several trigger words and one Dutch word each. Basic clause [5] contains only trigger words, and [6] only a Dutch word. Thus, there were no codeswitches in this fragment.

(20′) [1] *en ik moest naar Roermond*

[2] *want ik kon niet*

[3] *in Tegelen blijven*

[4] *daar is niets*

[5] *waar je kunt*

[6] *blijven*

At the clause level, the adjusted triggering theory presented by Broersma and De Bot (2006) predicts that codeswitching is more likely to take place in a basic clause that contains a trigger word than in a basic clause that does not contain a trigger word. To test this prediction, basic clauses with a trigger word were compared to basic clauses without a trigger word. Examples of basic clauses with a trigger word that contain a codeswitch are (18) and (19′) [3,9], and of basic clauses with a trigger word that do not contain a codeswitch are (19′) [2,5,6,7] and (20′) [1,2,3,4]. An example of a basic clause without a trigger word and without a codeswitch is (20′) [6]. Table 4 shows

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

Table 4. *Number of basic clausescontaining a codeswitch, number of basic clauses not containing a codeswitch, and percentage of basic clauses containing a codeswitch; split by basic clauses containing a trigger word and basic clauses not containing a trigger word.*

Trigger word in basic clause	Codeswitch		
	Yes	No	% Yes
Yes	104	173	37.5
No	2	27	6.9

Table 5. *Number of basic clauses that were entirely codeswitched, number of basic clauses not containing a codeswitch, and percentage of basic clauses that were entirely codeswitched; split by basic clauses containing a trigger word and basic clauses not containing a trigger word.*

Trigger word in basic clause	Codeswitch		
	Yes	No	% Yes
Yes	60	173	25.8
No	2	27	6.9

that 37.5% of the basic clauses with a trigger word contain a codeswitch, versus 6.9% of the basic clauses without a trigger word. This difference is statistically significant ($\chi^2 = 10.89$, $p < .001$, $P < .001$). Thus, codeswitching is indeed more likely to take place in a basic clause that contains a trigger word than in a basic clause that does not contain a trigger word.

Among the 104 basic clauses containing a trigger word and a codeswitch, there were 60 inter-clausal codeswitches, where the entire basic clause containing the trigger word was codeswitched. Examples are (19′) [9] and (21) [2].

(21) [1] AND THEY WERE SO SURPRISED

[2] *en toen van Heerlen ben ik weer terug gegaan al op de fiets en over de brug en zo naar Utrecht over de Vecht met een pont*

"And they were so surprised. And then from Heerlen I went back again, on my bike, and over the bridge and so to Utrecht over the Vecht by ferry."

Note that this analysis is different from the previous ones in that it implies that the language of an entire basic clause can be determined. For trigger words, it cannot be determined unambiguously which language they come from. Thus, the language of a basic clause is determined by its non-trigger words. Sometimes, clauses contain only few non-trigger words, as (19′) [9] does. For all instances of inter-clausal codeswitches, the phonological realization of the trigger words was found to be in line with the language of the non-trigger words in the basic clause. Narrowing the analysis down to inter-clausal codeswitches, Table 5 shows that 25.8% of the basic clauses with a trigger word were codeswitched, versus 6.9% of the basic clauses without a trigger word, which is a statistically significant difference ($\chi^2 = 5.08$, $p < .05$, $P < .05$).

3.3 Types of trigger words

So far in the analyses, all trigger words have been considered together. However, within the category of trigger words, several types of words can be distinguished. Whereas the Dutch – Moroccan Arabic data contained only 104 trigger words, the present data contained 2035 trigger words. Therefore, these data are very suitable for a further assessment of which types of words exactly can trigger a codeswitch.

First, there are proper nouns among the trigger words. In the Dutch – Moroccan Arabic data, the majority of the trigger words were proper nouns. In the present data, proper nouns form only a small subset of all the trigger words, and therefore their triggering potential can be compared to that of other trigger words. Second, a distinction can be made between content words and function words. Whereas in the Dutch – Moroccan Arabic data, all trigger words were nouns, in the present data, trigger words comprise both content words and function words. Third, trigger words were defined such that some variation in the Dutch and English phonological forms was allowed. Due to the large number of trigger words, the data lend themselves well to a further assessment of the importance of the amount of form overlap for triggering. There are enough trigger words to categorize them into sets with more and less form overlap and to investigate their triggering potential separately.

In order to investigate the triggering potential of different types of cognates, all trigger words in the data were divided into four categories. First, a distinction was made between content words and function words. Next, content words were further divided into proper nouns, other content words with good form overlap, and other content words with moderate form overlap. For each basic clause, it was first determined whether it contained a trigger word that was a proper noun or another content word with good form overlap, or whether it contained a trigger word that was a content word with moderate form overlap, or else whether it contained a trigger word that was a function word. As most basic clauses contained

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

Table 6. *Number of basic clauses containing a codeswitch, number of basic clauses not containing a codeswitch, and percentage of basic clauses containing a codeswitch; split by basic clauses containing different types of trigger words.*

Trigger word in basic clause	Codeswitch		
	Yes	No	% Yes
Proper noun	10	20	33.3
Content word, good form overlap	27	27	50.0
Content word, moderate form overlap	49	76	39.2
Function word	18	50	26.5
Total trigger words	**104**	**173**	**37.5**
None	2	27	6.9

several trigger words, only the trigger word that came first in this order was used to categorize the basic clauses. (For example, if a basic clause contained a content word with good form overlap, the presence of function-word triggers was not taken into account anymore.)

The analyses were based on the first judge's classification. To establish the reliability of this classification, all six judges categorized the trigger words. Inter-rater reliability was high again ($ICC = .825$, $p < .001$). The correlation between the first judge's coding and each of the five other judges' coding was medium (averaged Pearson r (2023.2) $= .5038$, $p < .01$), but note that it was higher than the correlation averaged over all pairs of judges (r (2023.2) $= .4625$, $p < .01$), showing that the medium correlation was not due to the first judge.

For example, all judges classified *Nieuw Zeeland* (/niwˈzeːlant/ "New Zealand") and "Ellen" (/ˈɛlən/) as proper noun triggers, *probleem* (/proːˈbleːm/ "problem"), *telefoon* (/teːləˈfoːn/ "telephone"), and GROUP (*groep* /χruːp/) as content words with good form overlap, *tijd* (/tɛit/ "time") and TIME as content words with moderate form overlap, and *in* (/ɪn/) – IN, and *de* (/də/) – THE as function-word triggers.

Table 6 shows the percentage of basic clauses containing a codeswitch for each category of trigger words. Basic clauses containing a proper noun contained a codeswitch 33.3% of the time, basic clauses containing a content word with good form overlap 50.0%, and basic clauses containing a content word with moderate form overlap 39.2% of the time. Basic clauses with only a function-word trigger contained a codeswitch 26.5% of the time. For each category of trigger words, significantly more codeswitches were found than if no trigger word was present (proper nouns: $\chi^2 = 6.36$, $p < .05$, $P < .05$; good overlap: $\chi^2 = 15.42$, $p < .0001$, $P < .0001$; moderate overlap: $\chi^2 = 11.09$, $p < .001$, $P < .001$; function words: $\chi^2 = 4.76$, $p < .05$, $P < .05$). Whereas there is some variation in the percentage of codeswitches, comparing the different types of trigger words with χ^2 tests did not yield any significant differences.

Thus, the data show that proper nouns trigger codeswitches in the same way as other cognates do, that both content-word and function-word cognates trigger codeswitches, and that moderate form overlap is enough for a cognate to function as a trigger word.

4. General discussion

The data show clear evidence of triggered codeswitching: codeswitches are more often found when there is a trigger word than when there is none. Dutch and English are related languages and share many cognates; in the present data, 71.4% of all words were categorized as trigger words. Apparently, the large proportion of trigger words did not reduce their potential to induce codeswitching.

Previous work (Broersma and De Bot, 2006) showed that in a Dutch – Moroccan Arabic data set where the cognates consisted mainly of proper nouns, these cognates triggered codeswitching. In the present data set, a large variety of cognates was present. It was found that proper nouns as well as other cognates could trigger codeswitching. Thus, the evidence found for the Dutch – Moroccan Arabic data was replicated and extended to other types of trigger words. Further, it was shown that cognates could trigger codeswitches if they had strongly overlapping forms in the two languages, but also if their form overlap was only moderate. Apparently, even with the large proportion of trigger words in the data, even the less well-matching cognates still had the potential to influence language choice. Finally, both content-word cognates and function-word cognates could trigger codeswitches. Thus, triggering is not limited to a particular type of cognates, but seems to be a general property of cognates. Therefore, it might be a common phenomenon in codeswitching situations.

Triggered codeswitching was investigated both at the word level and at the clause level. Of course, the larger scope of the clause level also implies making less specific predictions than can be made at the word level. Therefore, Broersma and De Bot (2006) determined whether the broader view of the clause level contributed substantially to the more specific predictions at the word level in explaining their Dutch – Moroccan Arabic data. A joint analysis of the two versions of the triggering theory showed that most of the codeswitches that the word-level view explained were also accounted for at the clause level, whereas the clause-level approach additionally explained many cases of codeswitching that the word-level view could not account for. Thus, the clause-level view covered a substantial amount of codeswitches that the word-level approach left unaccounted for. Unfortunately, a similar joint analysis is not possible with the present data. Due

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

to the pattern of rapid codeswitching (e.g., see (18), where the non-trigger words are consecutively English, Dutch, English, Dutch, and English), it is not clear which codeswitches are exclusively explained by either version of the triggering theory. The two versions follow a different definition of what counts as one codeswitch. In terms of the word-level approach, each word can be a codeswitch and therefore basis clauses can contain multiple codeswitches, whereas for the clause-level view, each basic clause can count as one codeswitch only. In a single basic clause, some codeswitches may be explained and others unexplained by the word-level approach. The question is then when the codeswitch explained by the clause-level approach should be considered to be explained by the word-level approach as well; if the word-level approach has to explain only one codeswitch within the basic clause, this disfavors the clause-level approach, and if the word-level approach has to explain all codeswitches within the basic clause, this disfavors the word-level approach. (Note that this was not a problem for the Dutch – Moroccan Arabic data described by Broersma and De Bot (2006), as those data did not show a similar pattern of multiple codeswitches within a basic clause.) Being maximally critical to the broader clause-level approach (such that the word-level approach has to explain only one codeswitch within the basic clause), there is only one codeswitch in the data that the word-level approach can, and the clause-level approach cannot, explain, and there are ten codeswitches that the clause-level approach can and the word-level approach cannot explain. Neither of these effects remain significant. Thus, it can be concluded that the word-level approach does a very good job explaining the triggered codeswitches in the data without the additional broader predictions at the clause level.

Note, however, that the clause level does help interpret the codeswitching pattern in the data, as it shows that codeswitches often occur between basic clauses rather than within basic clauses. In the present study, 58.5% of all the codeswitches in the data was inter-clausal (versus 41.5% intra-clausal), and of all the codeswitches that the clause-level approach explained, 56.6% was inter-clausal. Thus, when a trigger word triggered a codeswitch, in about half of the cases, the entire basic clause that contained the trigger word, rather than only part of it, was codeswitched. Similarly, in the Dutch – Moroccan Arabic data (Broersma and De Bot, 2006), 39.6% of all the codeswitches in the data and 47.6% of all the codeswitches that the clause-level approach explained was inter-clausal (i.e., an alternational codeswitch; Muysken, 2000).

Thus, triggering at the word level could explain the codeswitches in these data better than the codeswitches in the Dutch – Moroccan Arabic corpus. In the present data, codeswitches more often immediately bordered on a trigger word, whereas in the Dutch – Moroccan Arabic corpus, a larger proportion of the codeswitches could not be explained by an immediately neighboring trigger word, but only by a trigger word somewhere else in the basic clause. This might be due to the fact that there were more trigger words in the Dutch–English data than in the Dutch – Moroccan Arabic data, and to the fact that there was much more codeswitching in the former data (where 17.0% of the non-cognate words were codeswitched) than in the latter data (with 2.8%). In the Dutch – Moroccan Arabic speech, if a trigger word did not lead to a codeswitch immediately, its influence – increasing the activation of the language that was not being spoken – might have remained for a while and might have resulted in a codeswitch after some delay. In the Dutch–English speech, if one trigger word did not lead to a codeswitch immediately, another trigger word would soon occur, and this one might tip the balance and lead to a codeswitch. Thus, the trigger word that induced the codeswitch might have done so more quickly here than in the Dutch – Moroccan Arabic data, because other trigger words had already prepared the codeswitch. Another reason that the codeswitches occurred closer to the trigger words might be that it was easier to tip the balance in the Dutch–English speech, as more rapid codeswitching was taking place in the Dutch–English speech anyway. Thus, one trigger might be enough to immediately cause a codeswitch, with no delay.

An important difference between the two data sets is that in the Dutch – Moroccan Arabic data, words immediately following a trigger word had an increased chance of being codeswitched, but words immediately preceding one did not, whereas in the Dutch–English data, both words following and words preceding a trigger word had an increased chance of being codeswitched. As Broersma and De Bot (2006) point out, the order in which words are selected is not necessarily the order in which they end up in the surface structure of a sentence and in which they are eventually pronounced. Thus, a word that is selected from the mental lexicon immediately after a trigger word might end up before it in the sentence as it is pronounced. If in the Dutch – Moroccan Arabic data, as argued above, there was often some delay between a trigger word and the codeswitch it induced, the chance was smaller that the codeswitch ended up before the trigger word than in the Dutch–English data. Thus, triggered codeswitching of preceding words might only occur in dense codeswitching situations.

The evidence of triggered codeswitching between the typologically related languages Dutch and English is in line with the history of the triggering hypothesis. Clyne originally based this hypothesis on observations of the language use of German–English (1967, 1972) and Dutch–English (1977, 1980) bilingual immigrants in Australia. Thus, the first observations to be made about triggered codeswitching occurred within pairs of Germanic languages. Clyne (1980) already proposed, however, that triggered codeswitching might also occur

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

between unrelated languages like Estonian and English, and in later work he presented many examples of co-occurring cognates and codeswitches in language pairs with varying degrees of typological closeness (2003). These ideas are now statistically corroborated: in addition to the evidence that was previously found for triggered codeswitching between Dutch and Moroccan Arabic (Broersma and De Bot, 2006), which are unrelated languages with little lexical overlap, the present study provides evidence for triggered codeswitching between lexically related languages. These data, with their dense codeswitching pattern and large number of trigger words, provide a valuable extension of the earlier findings. They show that triggered codeswitching occurs between typologically related languages as well as typologically unrelated languages, but that there might be differences in the patterns of codeswitching, depending on the amount of trigger words and the amount of codeswitching in the bilinguals' speech.

Further, the results show that triggered codeswitching is not limited to a specific population but occurs for bilingual speakers from different populations and with different language histories, and that triggered codeswitching occurs in different settings: in Broersma and De Bot (2006) in conversation among peers, and in the present study in an interview with a researcher. Thus, triggered codeswitching might be a common phenomenon in bilingual speech. How common is something that future studies, with more speakers and different language pairs, need to establish.

References

Backus, A. (1996). *Two in one: Bilingual speech of Turkish immigrants in The Netherlands*. Tilburg: Tilburg University Press.

Backus, A. (2009). Codeswitching as one piece of the puzzle of language change: The case of Turkish *yapmak*. In Isurin et al. (eds.), pp. 307–336.

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*, pp. 407–434. New York: Holt, Rinehart and Winston.

Broersma, M. (2005). Perception of familiar contrasts in unfamiliar positions. *Journal of the Acoustical Society of America*, 117, 3890–3901.

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.

Broersma, M., Isurin, L., Bultena, S. & De Bot, K. (2009). Triggered codeswitching: Evidence from Dutch–English and Russian–English bilinguals. In Isurin et al. (eds.), pp. 103–128.

Bullock, B. E. & Toribio, A. J. (2009). Trying to hit a moving target: On the sociophonetics of codeswitching. In Isurin et al. (eds.), pp. 189–206.

Byrd, D. (1993). 54,000 American stops. *UCLA Working Papers*, 83, 97–115.

Clyne, M. (1967). *Transference and triggering: Observations on the language assimilation of postwar German-speaking migrants in Australia*. The Hague: Martinus Nijhoff.

Clyne, M. (1972). *Perspectives on language contact: Based on a study of German in Australia*. Melbourne: Hawthorn.

Clyne, M. (1977). Nieuw-Hollands or Double Dutch. *Dutch Studies*, 3, 1–20.

Clyne, M. (1980). Triggering and language processing. *Canadian Journal of Psychology*, 34, 400–406.

Clyne, M. (2003). *Dynamics of language contact: English and immigrant languages*. Cambridge: Cambridge University Press.

Cutler, A., McQueen, J. & Robinson, K. (1990). Elizabeth and John: Sound patterns of men's and women's names. *Journal of Linguistics*, 26, 471–482.

De Bot, K., Broersma, M. & Isurin, L. (2009). Sources of triggering in codeswitching. In Isurin et al. (eds.), pp. 85–102.

De Bot, K. & Clyne, M. (1989). Language reversion revisited. *Studies in Second Language Acquisition*, 11, 167–177.

De Groot, A. M. B. & Nas, G. L. J. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language*, 30, 90–123.

Desmet, T. & Declercq, M. (2006). Cross-linguistic priming of syntactic hierarchical configuration information. *Journal of Memory and Language*, 54, 610–632.

Dijkstra, T., Grainger, J. & Van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homophones: The neglected role of phonology. *Journal of Memory and Language*, 41, 496–518.

Dijkstra, T. & Van Hell, J. (2003). Testing the Language Mode hypothesis using trilinguals. *International Journal of Bilingual Education and Bilingualism*, 6, 2–12.

Dong, Y., Gui, S. & MacWhinney, B. (2005). Shared and separate meanings in the bilingual mental lexicon. *Bilingualism: Language and Cognition*, 8, 221–238.

Flege, J. E. & Eefting, W. (1987). Cross-language switching in stop consonant perception and production by Dutch speakers of English. *Speech Communication*, 6, 185–202.

Flege, J. E., Schirru, C. & MacKay, I. R. A. (2003). Interaction between the native and second language phonetic subsystems. *Speech Communication*, 40, 467–491.

Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, 1, 131–149.

Hartsuiker, R. J., Pickering, M. J. & Veltkamp, E. (2004). Is syntax separate or shared between languages? *Psychological Science*, 15, 409–414.

Hasselmo, N. (1961). *American Swedish*. Ph.D. dissertation, Harvard University.

Hulsen, M. (2000). *Language loss and language processing: Three generations of Dutch migrants in New Zealand*. Ph.D. dissertation, Nijmegen University.

Isurin, L., Winford, D. & De Bot, K. (eds.) (2009). *Cross-disciplinary approaches to codeswitching*. Amsterdam: John Benjamins.

Kroll, J. F., Bobb, S. C. & Wodniecka, Z. (2006). Language selectivity is the exception, not the rule: Arguments

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press

against a fixed locus of language selection in bilingual speech. *Bilingualism: Language and Cognition*, 9, 119–135.

Lemhöfer, K. & Dijkstra, T. (2004). Recognizing cognates and interlingual homographs: Effects of code similarity in language-specific and generalized lexical decision. *Memory & Cognition*, 32, 533–550.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.

Muysken, P. (2000). *Bilingual speech: A typology of code-mixing*. Cambridge: Cambridge University Press.

Myers-Scotton, C. (1993). Common and uncommon ground: Social and structural factors in codeswitching. *Language in Society*, 22, 475–503.

Myers-Scotton, C. (1997). *Duelling languages: Grammatical structure in codeswitching*. Oxford: Oxford University Press.

Myers-Scotton, C. (2006). Natural language knocks on the laboratory door. *Bilingualism: Language and Cognition*, 9, 203–212.

Odlin, T. (1989). *Language transfer*. Cambridge: Cambridge University Press.

Odlin, T. (2009). Transfer and code-switching: Separate territories but common concerns on the border. In Isurin. et al. (eds.), pp. 337–358.

Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.

Roelofs, A. & Verhoeff, K. (2006). Modeling the control of phonological encoding in bilingual speakers. *Bilingualism: Language and Cognition*, 9, 167–176.

Schulpen, B., Dijkstra, T., Schriefers, H. J. & Hasper, M. (2003). Recognition of interlingual homophones in bilingual auditory word recognition. *Journal of Experimental Psychology: Human Perception and Performance*, 29, 1155–1178.

Van Hell, J. & De Groot, A. M. B. (1998). Conceptual representation in bilingual memory: Effects of concreteness and cognate status in word association. *Bilingualism: Language and Cognition*, 1, 193–211.

Van Hell, J. & Dijkstra, T. (2002). Foreign language knowledge can influence native language performance in exclusively native contexts. *Psychonomic Bulletin & Review*, 9, 780–789.

Van Lancker, D. & Klein, K. (1990). *Preserved recognition of familiar personal names in global aphasia*. Presented at the Eighteenth Annual Meeting of the International Neuropsychological Society, Orlando, FL.

Winford, D. (2009). On the unity of contact phenomena and their underlying mechanisms: The case of borrowing. In Isurin. et al. (eds.), pp. 279–305.

Received: March 26, 2008
Final revision received: August 15, 2008
Accepted: August 30, 2008

https://doi.org/10.1017/S1366728909990204 Published online by Cambridge University Press