
Radboud Universiteit

# Code-switching by Spanish/English bilinguals in online reviews

Bachelor Thesis

Student: Juul Meijerink

Supervisor: Dr. S.N. Barasa

International Business Communication

Faculty of Arts

Radboud University

05-06-2021

**Abstract**

The present corpus study aimed to obtain an improved understanding of why and how code-switching occurs in written form. This study focused on online reviews written by Spanish/English bilinguals. Specifically, this study examined if the Type of code-switching in reviews was based on the Type of review (like/dislike/neutral) and what the main motivation to code-switch was. Furthermore, it was hypothesized that Spanish/English bilinguals are more likely to code-switch to English in neutral reviews due to less emotional involvement. Most studies and theories on code-switching concern spoken code-switching, leaving a gap in research, while the analysis and understanding of code-switching in e.g. online reviews could also be of value for businesses. The corpus consisted of 92 different hotel reviews, of which the code-switches were coded for Type of code-switch (inter/intra-sentential, tagging and word-affixation), Type of review (like/dislike/neutral) and Motivation to CS by two different coders. Statistical tests showed no significant relationship for Type of code-switch and Type of review. The most frequent motivations were Anglicism, Exhibitionism and Emphasis on message. The hypothesis was refuted, as emotion did not seem to influence CS in online reviews. The results were partly in contrast with earlier found theories on code-switching, the explanation mainly being the influence of the internet and certain possible limitations in the corpus sample.

*Keywords*: Online reviews, code-switching, Spanish/English bilinguals, emotion

## Introduction

The world is overflowing with people, with many different cultures, values, traditions and languages. Whether they like it or not, humans are bound to communicate with each other in order to reside on earth. The act of communicating with each other is largely performed through language. Taking into consideration that so many different people are trying to communicate with one another, it is perhaps not unexpected that many master two or more different languages. Generally, besides monolingualism - when one is fluent in one language -, a distinction may be made between someone who is fluent in two languages, who would be called bilingual and someone who is fluent in more than two languages, who would be called multilingual. Bilingualism seems to be an emergent research topic when considering the number of recent studies (e.g. Altarriba & Heredia, 2018; García & Wei, 2014; Kroll, Dussias, Bice & Perroti, 2015) and the globalisation of the last decade probably partly affected this. However, bilingualism is ongoing and has been ongoing for a long time. The world was simply too preoccupied with monolingualism to take notice and interest in it.

Bilingualism, at times, can lead to a linguistic phenomenon known as code-switching, often appearing in speech acts, which are utterances that serve a function in communication, varying from one word to several sentences (CARLA, 2021). Broadly, code-switching can be defined as the substitute use of two different languages in one speech act (Auer, 2013; Mabule, 2015) albeit that there is no generally agreed definition. Nonetheless, switching between three or more codes is likewise classified as code-switching. Myers-Scotton (1993a) and Nilep (2006), for instance, define it as the alternation of linguistic elements or in other words, contextualization of talk in speech acts. This study will work with the broad definition of code-switching while focusing on code-switching by Spanish/English bilinguals.

The study of code-switching is relevant for research as it incorporates the link between language use in a social setting and the linguistic form, two variables that were long seen as separate instead of interrelated (Heller & Pfaff, 2008). Additionally, code-switching can be found in both real-life conversations as in written discourse. Nevertheless, there has been an unevenly distributed interest in the different displays of code-switching (henceforth: CS). As Sebba, Mahootian and Jonsson (2012) also concluded; even though an increase can be seen in studies on bi- and multilingualism and CS, which is a great step forward for the in-depth understanding of the subject matter, said studies mainly focus on spoken discourse (e.g. Adendorff, 1993; Auer, 2013; Grim, 2008). What remains is a great gap of knowledge

on written discourse, specifically on online communication (Callahan, 2004; Weston & Gardner-Chloros, 2015). Moreover, broader research into CS by bilinguals might uncover more on the process of bilingual communication, which in turn might shed light on the functioning of the bilingual mind in general, of which supposedly only the surface has been scratched (Bhatia & Ritchie, 2004; Bialystok, Craik & Luk, 2012). This study aims to obtain an improved understanding of why and how CS occurs in written form, and in this, it focuses on Spanish/English bilinguals who CS in online reviews.

The study of CS can be approached in three different ways. The structural approach is directed at CS in terms of the language structure (e.g. syntax or lexicon); the psycholinguistic approach focuses on the underlying nature of the production and perception of CS and bilingualism. The sociolinguistic approach is concerned with the social factors that may or may not have an impact on CS (Stell & Yakpo, 2015). The subsequent theories will mainly fall under the sociolinguistic perspective, while later on psycholinguistic findings on CS will be explained. The structural approach focuses mainly on finding internal grammatical constraints at the occurrence of CS, which is not of interest to the current study and therefore not analysed in depth.

Moreover, Hoffmann (1991) states that there are distinct types of code switches. There can be intra-sentential CS, where the code is switched within a sentence or speech act. A fine example is the title of Poplack's (1980) study "*Sometimes I'll start a sentence in Spanish y termino en Español*". This can be seen as the most fluency demanding type of CS as the syntax of the sentence must be correct in both languages. Then there is inter-sentential CS, where code is switched between sentences. This type requires some fluency as both sentences should follow syntactical rules of corresponding languages. The third CS type is called tag-switching or extra-sentential, with insertion of tag-phrases or words in another language (e.g. *it's always like that, verdad?*). A last possible type of CS is word affixation, for instance adding an English suffix like -ing to a Spanish verb in order to create a new word or meaning.

In the first place, CS does not always happen at the appropriate place or time as Weinreich (1953) suggests; therefore, there exist many different theories that try to explain why and when people CS. However, most theories are again, based on spoken CS rather than written. Auer (2013) and Wei (1998) presume that to make sense of any CS theory, the initial focus should lie on how CS occurs. Following their theory of conversation analysis (CA), the

meaning of CS should be interpreted in relation to the speech sequences. This would mean that the occurrence of CS is related to what has been said before in the conversation, which is somewhat connected with the markedness model and Gumperz' distinction, which will be explained shortly. Nonetheless, the applicability of this CA theory depends on which approach to CS the study is aimed at.

Initially, Weinreich was one of the first to start defining and describing code-switching. Weinreich (1953) sees spontaneous CS of bilinguals within the same situation or sentence as flawed since bilinguals should possess the languages as two separate varieties for separate occasions. The alternation in inappropriate places or times might be the result of poor education and is believed to be random. Other scholars see CS as a deliberate activity, with different functions and intentions behind it (Blom & Gumperz, 1972; Gumperz, 1976; Myers-Scotton, 1983; 1993a). Blom & Gumperz (1972) for instance, made a distinction between situational and metaphorical CS. Situational CS occurs when the bilingual speaker code-switches in order to redefine a situation because the speaker may be stimulated by the conversation topic, setting or fellow interlocutor (Grim, 2008). Blom & Gumperz (1972, p.408) identify it as metaphorical switching when CS "enriches a situation, allowing for allusion to more than one social relationship within the situation". This would mean that one would switch from one language variety to the other in the same situation, for instance, to discuss work at home.

Later on in 1982, Gumperz acknowledged that this distinction was quite hard to make when analysing CS; resulting in a new term, conversational CS, which can be defined as the juxtaposition from two different grammatical systems (languages) within the same speech exchange (Nilep, 2006). To still be able to identify the function of CS, Gumperz (1982) identified six discourse functions of CS in speech exchange to epitomize its prevalent use: quotations, addressee specification, interjections, reiteration, message qualification and personalization versus objectivization. Although this semantic model accounts to some extent for why CS occurs, it can be seen as deficient, as not all the functions reveal what the person achieves through the CS for instance (Boztepe, 2003).

Myers-Scotton, who is quite renown in the field of sociolinguistics, opposes Gumperz' theories as he does not attempt to explain the general presence of CS (1993a). In turn, Myers-Scotton tries to offer an overall theoretical explanation of the sociolinguistic and pragmatic aspects of CS, by means of the Markedness-model. The theory behind this model

suggests that people try to redefine interaction and social meaning through CS (Myers-Scotton & Ury, 1977). The model itself is mainly based on code decisions that are either unmarked or marked. The unmarked choice of switching is equivalent to the expected choice of switching; within this CS choice the speaker switches from one expected code to another because of the change in situation. The cohesion between code and situation involves what Myers-Scotton (1998) calls: the rights and obligations (RO) set, which incorporates all associations of the behavioural norm within a social group. Because of this, certain choices would be perceived as either marked or unmarked; a code choice aligned with the social groups' norm prediction would thus be unmarked. A marked choice, in turn, distances itself from the existing or expected RO set in order to create a new set to, for instance, increase the (social) distance between interlocutors.

These marked CS choices regularly carry some sort of meta-message within (Kieswetter, 1995), as they intend to change aspects of the interaction. Sometimes what happens is that interlocutors are not sure of the expected RO set in certain situations, this might lead to interlocutors not choosing one specific code in order to stay on the middle ground until they find the successful code, which is called exploratory CS (Myers-Scotton, 1993a). This model covers quite some theoretical ground on why and how people make the choice to CS, however, it also suggests that CS comes from an entirely rational incentive, opposing the belief that CS sometimes happens unconsciously or at least not strategically (Woolard, 2004). Another weakness of this model is that it is mainly relying on external intelligence; analysts make assumptions about interlocutors' beliefs instead of using facts (Auer, 2013).

As Gardner-Chloros & Weston (2015) point out, many of the sociolinguistic models are based on spoken discourse and it is still unclear to what extent these models can be applied to written discourse and text. Subsequently, after checking to see if there exists an overlap in the functions of CS on written and spoken text, it was concluded that there is a limited overlap. Myers-Scotton's (1993b) Matrix Language Frame Model (MLF), in particular, worked moderately well when analysing literature. This model suggests that in a bilingual person's speech, one of the two languages bears a dominant role in the grammar of the utterance. The dominant language is what we call the matrix language, and the other language, the embedded language. The theory behind this model suggests that during CS, there still is an uneven relationship between languages, as the matrix language offers the

morphosyntactic frame of the sentence while the embedded language plays a supplementary role (Myers-Scotton, 1993b). A disadvantage of the MLF model is that it barely explains one type of CS, nor does it cover the issue of why CS occurs.

In line with the abovementioned theories, it is at least suspected that CS is occasionally a deliberate course of action. Many scholars have tried to identify why bilingual speakers code-switch, what the intrinsic or extrinsic motivation behind it is. Results vary widely, as bilinguals may code-switch to express solidarity with a certain group of people (Gal, 1978), to repeat a message and emphasize a point (Gal, 1979; Gumperz, 1982), to create a more effective production of words (Simin & Hamid, 1994) or to signal a change of topic in the conversation (Fishman, 1972; Hoffman, 1991). A more elaborated motivation is to use CS as a distancing function. Switching code could allow the bilingual to speak about topics that they would find too disturbing or upsetting in their first language (Bond & Lai, 1986; Camilleri, 1996; Ladegaard, 2018). Myers-Scotton (1993a) briefly describes motivations to make marked choices in her theory of the Markedness model: inter alia, to increase social distance via authority, to emphasise the message, to communicate irony and for aesthetic effect.

These motivations to CS, however, mostly appeared in spoken discourse; few scholars have conducted a study on written discourse, let alone, online communication. Barasa (2010) is one of the scholars that did, in a study of what back then was still called Computer Mediated Communication. The following additional reasons to CS were found: least effort principle, to write a message with as little time or effort as possible ; rapidity, the code-switched word is the first one that pops up in the user's mind; creativity; mode limitation, CS to ensure brevity; to show identity; enhance accuracy, to choose the word in the language that is most adequate (Barasa, 2010). Additionally, as an extension to Barasa's findings, Gammaldi (2016) included a few categories based on her studies' observations, such as: self-censorship, exhibitionism and Anglicism.

Alternatively, there are studies such as Wardhaugh (1998) that claim that CS at times happens subconsciously. This would mean that bilinguals do not always rationally make the choice to switch codes, thus there must be another factor. A factor that potentially has an influence on CS, is emotion. Several studies demonstrate that the emotional state of a person might influence someone's language preference (Pavlenko, 2002). The preferred language for communicating strong emotions such as anger appears to be the speaker's first language (L1).

Moreover, swear words for instance were felt to have greater emotional force in L1 than in L2 (Dewaele, 2004). Allegedly, bilinguals even feel different when talking in their L1 compared to their L2, as they reported feeling more authentic (less fake), more logical, more serious and more emotional in their L1 (Dewaele, 2006; Dewaele & Nakano, 2013). Nevertheless, in practice, this could lead to bilinguals CS to their L2 due to the aforementioned distancing motivation. Other studies show that when bilinguals have only achieved fluency in L2 after their childhood, there might be a difference in emotional impact, which could manifest as following: L1 being the language in which to express personal involvement and L2 being the language of distancing and detachment (Bond & Lai, 1986; Amati-Mehler, Argentieri, Canestri & Whitelaw-Cucco, 1993; Anooshian & Hertel, 1994). Specifically, bilingual speakers may switch to their

6
6

This current study will focus on a specific form of communication known as online communication appearing as online product/service reviews. One can communicate to others by writing a negative, positive or neutral review. These online evaluations can be of great importance to organizations as it influences consumer behaviour since it functions as "digital word of mouth" (Bonner and Willemsen, 2012). It would be interesting to see if already found motivations for CS in spoken discourse and Barasa's CMC apply for online reviews as well, seeing that online reviews differ in terms of situation and interaction time.

In short, there is a need for a better understanding of why CS occurs in online review writing by Spanish/English bilinguals and if this has any relation with the type of review that is being written. More specifically, the following research questions need to be addressed:

1. Is the type of CS in reviews based on the type of review (like/dislike/neutral)?
2. What is the main motivation for CS in Spanish/English bilinguals' reviews?

Since there is allegedly limited overlap with the already established CS functions of spoken text compared to written text (Gardner-Chloros & Weston, 2015), this research area is in need of new endeavours to create more extensive and accurate models than, for instance, the MLF model.

Apart from that, CS has occasionally been related to emotion, as the emotional state of a person could have an influence on the language preference (read: L1 or L2) (Pavlenko,

2002; Dewaele, 2006). Following Pavlenko's (2005) line of reasoning for this study, bilinguals might then be suspected to switch to their L1 when expressing more intense emotions and to their L2 when expressing less intense emotions. Thus, will emotions have an influence on the language preference in online communication and are bilinguals then likely to use L1 when writing an angry review for instance? Simultaneously, would that mean that when there is less emotional involvement when writing a review, Spanish/English bilinguals will express themselves regularly in their L2? This reasoning paved the way for the following hypothesis:

H1: Spanish/English bilinguals are more likely to CS to English in neutral reviews due to less emotional involvement.

This research will be a good start to reducing the scientific gap on written CS as well as providing more insight into the bilingual language culture. Furthermore, it might provide companies with a better understanding of the reviews given on their services and/or products. This again could be beneficial for the company's reputation and problem-solving in the long run, considering the correct interpretation of reviews could be of great value for organizations to find out how their service or product is experienced. It could help review analysts, for instance, to interpret the reviews at a more profound level when they understand the presence of CS alongside motivation and rating.

## Method

### *Materials*

This corpus study included data from the Spanish site Tripadvisor.es on which people voluntarily post reviews. In said reviews, people share their experience with a certain product or service. The Spanish TripAdvisor site has been chosen because this study focuses on Spanish/English bilinguals. Only reviews of hotels have been collected, so that the corpus data did not differ too much. Otherwise, the different topics of reviewed products or services might have played a role in the results of this study. As TripAdvisor only launched in 2000 on the internet, all reviews of the last 21 years have been considered as eligible to our study, provided that they had been posted online on the website and were in written form. The material selection of this study has not been completely random as only reviews that included one or more of the 4 types of CS (intra/inter-sentential, tag-switching and word affixation) were incorporated. Moreover, only reviews that included the Spanish and English language or the alternation of these two have been considered. That being said, any review containing these prerequisites had an equal chance of being incorporated.

In total, 92 reviews with Spanish/English CS were used to build the corpus, which contained 138 code-switches altogether. A minimum of 75 reviews had been set for the collection of the data, however, the three data collectors performing this task came up with a combined total of 92 reviews. All data was assembled in one SPSS file. Three exemplary reviews from the corpus can be found in Appendix B. The analysed units consisted of full reviews, which differentiated from a few words to a whole paragraph. Additionally, the star ratings above the written text of the reviews have been taken into account in order to be able to classify the reviews in the different types (like/dislike/neutral).

### *Procedure*

This study wished to analyse CS in written text. To achieve that, several variables were operationalised and then analysed. The variable ‘type of CS’ was operationalised through the four different categories of CS, employing the previously by Hoffman (1991) explained definitions of the type of code-switches. The variable ‘type of review’ was operationalised through the distinction of the emotional value of the review. This distinction implied that the reviews were either found to be positive, negative or neutral. The third variable is ‘type of

motivation', which was operationalised with the help of pre-existing categories in which each case of CS might fit, according to several theories explained later in the procedure.

Firstly, the reviews were considered as reviews containing CS, when the reviewer made substitute or alternating use of the languages Spanish and English within the borders of one review. The alternations have been taken into account in both ways (English language as matrix language and CS to Spanish and/or vice versa). All four types of CS have been perceived as CS; therefore, a review could contain either intra-sentential, inter-sentential, tag-switching or word affixation or a mix of these types. Subsequently, all marked CS were coded as either one of the above types of CS in the data file. In other words, all 138 code-switches received a coding score of either 1 (inter-sentential), 2 ( intra-sentential), 3 (Tagging) or 4 (word-affixation).

Secondly, the type of review was coded, this variable could be coded as like, dislike or neutral. All reviews have been evaluated through the rating by stars, ranging from 1 star to 5 stars. A review with 1-2 stars has been coded as 'dislike', a review with 3 stars as 'neutral' and a review with 4-5 stars as 'like'. Only one coder coded this variable, seeing that the number of stars that a review had received is quite straightforward and objective.

Thirdly, the motivation to CS was coded with help of already existing motivational theories such as Barasa's (2010), Myers-Scotton's (1993a) and Gammaldi's (2016), from which a combination of categories originated that seemed most suitable for this corpus. The reviews containing CS have been categorized by the coders based on the context of the code switch. Possible categories were: Least effort principle, Mode limitation, Accuracy, Emphasis on message, Communicating irony, Distancing function, Identity exposure, Exhibitionism, Anglicism and an 'other' category for switches that can not be placed under any of the above.

Furthermore, the corpus was coded by 2 different coders who are fluent or at least semi-fluent in both Spanish and English so that coding would not be compromised by inability of understanding. Both coders received additional explanation on the interpretation of the different motivational categories (found in Appendix C). The inter-rater reliability between the two coders of the variable "type of CS" was more than acceptable: $\kappa = 0.79$, $p < .001$. The inter-rater reliability of the variable "type of motivation" was also acceptable: $\kappa = 0.69$, $p < .001$. Finally, the two coders came together to discuss the items until they agreed on a final coding for each CS for both Type of CS as Motivation to CS. The final coding scores were then used in the statistical analyses.

***Statistical analysis***

A chi-square test has been done to determine if there is a significant relationship between the variables; type of CS and type of review (like/dislike/neutral). In this chi-square, the variable Primary First was used when analysing the data. This test was restricted to the use of Primary first so that no duplicate cases of the Type of Review were included in the calculation process. This was necessary since the 92 reviews contained 138 CS altogether, therefore, the data set would have calculated the review rating 138 times instead of 92 times. This could have influenced the results and thus the variable Primary First was switched on. Furthermore, descriptive statistics were used to test the main motivation to CS. Lastly, results from the first Chi-square have been used to test H1.

**Results**

This study aims to obtain an improved understanding of why and how CS occurs in written form, and in this, it focuses on Spanish/English bilinguals who CS in online reviews. A corpus that consisted of 138 code-switches in 92 reviews was used. With this corpus, the study aspired to test respectively the relation between the type of CS in reviews and the type of rating a review received (RQ1) and the main motivation for the said use of CS (RQ2). Furthermore, the study´s H1 tested if the bilinguals in this corpus were more likely to CS in neutral reviews as opposed to positive and negative reviews.

First of all, A Chi-square test was conducted to examine the relation between the type of CS (intra/inter-sentential, tag-switching and word affixation) and the type of review (dislike/neutral/like) (RQ1). This Chi-square showed no significant relation between type of CS and type of review ($\chi^2(4) = 3.85$, $p = 0.427$) . The results do show that most reviews were Intra-sentential, namely 73 (79.3%) of the 92 (100%) reviews. Plus, 55 (75.3%) of these 73 Intra-sentential reviews were positive. Additional exact measured results can be found in Table A.1 in Appendix A.

Secondly, this study investigated what the main motivation is to code-switch in the Spanish/English bilingual reviews of the corpus (RQ2). As aforementioned in the method section, the motivation was coded by means of the 10 different categories, which resulted in the following descriptive statistics. The frequencies and percentages of the different types of motivations to switch code can be found in Table 1. The main motivations for switching code

in this corpus study were Anglicism (29.7%) and Exhibitionism (28.3%) followed by Emphasis on message (21.0%). Remarkable is that the categories Distancing function (0.0 %) and Identity Exposure (0.0%) both did not once occur in accordance with the coding.

Table 1. Frequencies and percentages of the different type of motivations to CS *(1- Least effort 2- Mode limitation, 3- Accuracy, 4- Emphasis on message, 5- Communicating Irony, 6- Distancing function, 7- Identity exposure, 8- Exhibitionism, 9- Anglicism, 10- Other).*

| Type of motivation | Frequency (N) | Percent (%) |
|---|---|---|
| 1. Least Effort | 4 | 2.9 |
| 2. Mode limitation | 1 | .7 |
| 3. Accuracy | 17 | 12.3 |
| 4. Emphasis on message | 29 | 21.0 |
| 5. Communicating Irony | 3 | 2.2 |
| 6. Distancing function | 0 | 0.0 |
| 7. Identity exposure | 0 | 0.0 |
| 8. Exhibitionism | 39 | 28.3 |
| 9. Anglicism | 41 | 29.7 |
| 10. Other | 4 | 2.9 |
| Total | 138 | 100.0 |

Finally, to test H1, to see if Spanish/English bilinguals are indeed more likely to CS in neutral reviews due to less emotional involvement, the results of the Chi-square of RQ1 were looked at anew. The results showed to be inconsistent with the proposed hypothesis. As shown in figure 1, CS occurred most often in positive reviews (coded as like) with 79.3%, while the neutral reviews with CS only accounted for 10.9% and the negative reviews (coded as dislike) for 9.8%. According to the statistical output, there exist no significant differences between the types of ratings in all three categories. The scores per category show that the proportions do not differ significantly from each other at the .05 significance level, as can be seen in Table A.1.

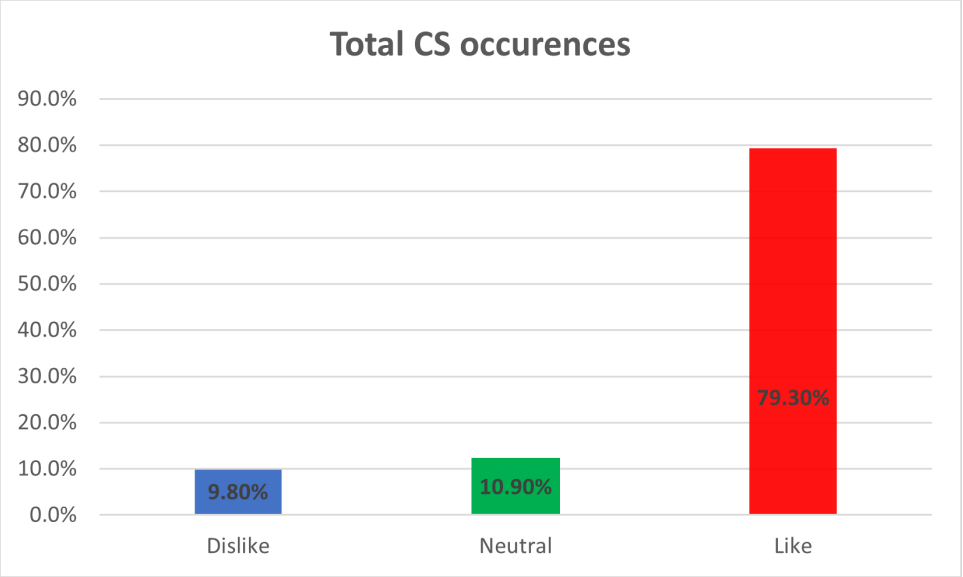



Figure 1: Total number (%) of CS occurrences per type of review.

**Conclusion and discussion**

Throughout this study, it was tested whether there exists a relation between the type of CS and the type of review (RQ1). However, in this study, no significant relationship between the type of CS and the type of review was found. In other words, Spanish/English bilinguals did not use intra-sentential CS more often when writing a negative review, nor did the other variables differ significantly or show any relation between each other, as can be interpreted from Table A.1. Still, it can be concluded that intra-sentential CS is the most used type of CS in this corpus, both for positive reviews and the reviews in general. This is remarkable as it is the most fluency demanding type of CS. Fluency could then possibly have an impact on the type of CS a bilingual is more likely to use, dividing the more and less fluent bilinguals from each other. This variable also reveals a limitation of this study since the corpus has limited to no factual data on the bilinguals. For instance, there is no assurance to the fact that the presumed bilinguals are actually fully bilingual, this is solely based on assumption. Nor is the degree of bilingualism checked, for this reason, nothing can be really

surmised from the possible variable of fluency. Thus, future research could benefit from including such factual data from the sample's reviewers.

Subsequently, the main motivation to code-switch was investigated (RQ2). As shown in the result section (see Table 1), the most frequent motivations were Anglicism, Exhibitionism and Emphasis on message. These motivations overlap somewhat with the motivations found in earlier studies on spoken discourse. Anglicism and Exhibitionism match with Gammaldi's (2016) findings, while Emphasis on message was previously stated as a preeminent category in the studies of Gal (1979), Gumperz (1982) and Myers-Scotton (1993a). The high score of Exhibitionism (see Appendix C for term explanation and interpretation) might be explained through the fact that the reviews on TripAdvisor have a broad and worldly public, perhaps reviewers feel more pressured to show off and create their bilingual identity online (Biró, 2019). The high frequency of CS motivation Emphasis on message could be explained by the fact that everyone in the world with internet access could fairly easily write an online review. Therefore, catching the attention amidst numerous other reviews might be believed to be achieved through the use of CS.

Nevertheless, CS as Distancing function and Identity Exposure were the lowest scoring categories and therefore on the contrary, not in accordance with previous studies. These current results seem to contradict the beliefs of Bond & Lai (1986) that CS as a distancing function is one of the main motivations as it did not once occur in the whole corpus and certainly not significantly more often for negative/dislike reviews. An explanation for these results might be found in a limitation of the present study, as the sample contained much more positive reviews compared to negative or neutral reviews (see Table 1). In other words, this sample can not confidently refute Bond & Lai's theory. To test these theories, future research should investigate the main motivation to CS in a sample with equal numbers of the different types of reviews.

Moreover, the conclusions of the study's two research questions (RQ1 + RQ2) seem to oppose Woolard's (2004) beliefs that CS happens unconsciously and non-strategically and seem more in conformity with Myers-Scotton (1993a) beliefs that CS is often a rational choice as the categories Exhibitionism and Emphasis on message for instance, turned out to be the more frequent motivations to CS. Whilst Anglicism seems to be more of an elementary motivation to CS, regarding the global use of English as Lingua Franca, the two previously stated categories could be seen as strategic CS. English as a global language and the internet,

in particular, seem to have a form of interaction, as many speakers come into contact via the internet and choose to speak English among each other (Herring, 2008). This could mean that the bilinguals might be more inclined to CS with Anglicism in mind because they are aware that they are communicating on the internet. However, just as Auer (2013) concluded regarding Myers-Scotton's (1993a) study, the current study is limited in the same form. Again, the coders made assumptions on the reviewers' motivations and it was, therefore, not based on factual information. Future research should try to avoid this, possibly by applying a different kind of research method to directly receive information from the reviewers. For instance, by sending out a follow-up questionnaire.

Thirdly, with regard to the H1, it is concluded that the Spanish/English bilinguals are not more likely to CS to English in neutral reviews due to less emotional involvement. The hypothesis is, therefore, refuted. These findings thus contradict Pavlenko's (2005) expectations that bilinguals CS to their L2 when expressing less intense emotions and L1 for more intense emotions. An explanation for these results might be that there might be in fact a difference between spoken and written CS interaction, seeing that when bilinguals write an online review it is not as prompt as in a face-to-face speech act, which might affect emotional experience and expression. It could also be that emotion does not influence CS sufficiently to cause an effect. This could be the result of a limitation in the current study, or it might mean that in the study of CS, research must look at different factors than emotion. Nevertheless, general conclusions can not be made based on this study, although it can not be ruled out that emotion is a connected variable to CS. It can only be concluded that in this study emotions did probably not affect CS that much.

A substantial remark has to be made concerning the hypothesis testing. There is reasonable doubt if there is anything to be concluded from this test since 73 out of 92 reviews are positive in comparison with 9 negative and 10 neutral reviews. There is a considerable chance that this distribution is not representative of all reviews online. Thus, this sample seems skewed and too narrow to draw any conclusions. The indicated is a limitation to RQ1 as well. This perceived flaw is presumably a result of the selection process. In the future, collecting a more equal number of negative/neutral/positive reviews will prevent this, so that some conclusions can be drawn about the division of CS in the reviews.

On the other hand, it is striking that there are so many more positive reviews with CS in this sample since the only requirement was that some form of CS needed to be present. The

review rating was not looked at during the selection process. Hence, this is interesting to investigate more extensively, to see if there is perhaps more frequent use of CS in positive reviews and what the reason for this could be.

One general limitation remained, as mentioned in the method (see statistical analysis) only the first code-switches of all 92 reviews were counted. Some reviews had more and different types of CS, these were not analysed. This may have influenced the outcome. Future research should take notice of this. Generally, future corpus studies could benefit from incorporating a larger corpus with both reviews with and without CS, to check relatively if there is a relation between the presence of CS and type of review. This recommendation and the previously mentioned may lead to a clearer and more factual understanding of CS and the motivation behind CS in online reviews.

To conclude, this study contributes to the still poorly researched research topic written CS. It has provided insight into possible informative variables for future research in both this field of research as in bilingual communication research. Although the tested variables did not show any significant relation, this study still provides relevant information about how CS works. Additionally, this study gained, to a certain extent, insight into how online reviews work for bilinguals. These insights can have practical implications for business analyses of reviews, for example, at the behest of hotel chains.

**References:**


Adendorff, R. (1993). Codeswitching amongst zulu‐speaking teachers and their pupils: Its functions and implications for teacher education. *Language and Education*, *7*(3), 141 - 162.

Altarriba, J., & Heredia, R. R. (Eds.). (2018). *An introduction to bilingualism: Principles and processes*. Routledge.

Anooshian, L. J., & Hertel, P. T. (1994). Emotionality in free recall: Language specificity in bilingual memory.*Cognition & Emotion*, *8*(6), 503-514.

Amati-Mehler, J., Argentieri, S., Canestri, J., & Whitelaw-Cucco, J. T. (1993). *The Babel of the unconscious: Mother tongue and foreign languages in the psychoanalytic dimension*. International Universities Press, Inc.

Auer, P. (Ed.). (2013). *Code-switching in conversation: Language, interaction and identity*. Routledge.

Barasa, S. (2010). *Language, mobile phones and internet: A study of SMS texting, email, IM and SNS chats in computer mediated communication (CMC) in Kenya.* Utrecht: LOT.

Bhatia, T.K., & Ritchie, W.C. (2004). Bilingualism in the Global Media and Advertising. In Bhatia, T.K & Ritchie, W.C. (2006*), Handbook of Bilingualism*, Chapter 20, pp. 512-546. Oxford: Blackwell Publishers Inc.

Bialystok, E., Craik, F.I., & Luk, G. (2012). Bilingualism: consequences for the mind and brain. *Trends in cognitive sciences, 16* (4), pp. 240-250.

Biró, E. (2019). Linguistic Identities in the Digital Space. *Acta Universitatis Sapientiae, Philologica*, *11*(2), 37-53.

Blom, J. P. & Gumperz, J. J. (1972). Social meaning in linguistic structures: Codeswitching in Norway. In J. J. Gumperz & D. Hymes (eds.). *Directions in Sociolinguistics* (pp. 407-434). New York: Holt, Rinehart & Winston.

Bond, M. H., & Lai, T. M. (1986). Embarrassment and code-switching into a second language. *Journal of Social Psychology, 126*(2), 179.

Bonner, F. & Willemsen, L. (2012). E-vloed op consumentengedrag. In Sak Boom, van den, Edith Smit & Suzanne De Bakker (Eds.), *Nachtmerrie of Droom. De ROI van customer media.* Relatiemedia deel 3 (pp. 125-136). Amsterdam: Customer Media Council.

Boztepe, E. (2003). Issues in code-switching: competing theories and models. *Studies in Applied Linguistics and TESOL*, *3*(2).

Callahan, L. (2004). *Spanish/English codeswitching in a written corpus* (Vol. 27). John Benjamins Publishing.

Camilleri, A. (1996). Language values and identities: Code switching in secondary classrooms in Malta. *Linguistics and education*, *8*(1), 85-103.

CARLA – Center for Advanced Research on Language Acquisition. What is a speech act?, Accessed 15/03/2021 from http://www.carla.umn.edu/speechacts/definition.html.

Dewaele, J. M. (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of multilingual and multicultural development*, *25*(2-3), 204 - 222.

Dewaele, J. (2006). Expressing anger in multiple languages. *Bilingual education and Bilingualism*, *56*, 118.

Dewaele, J. M., & Nakano, S. (2013). Multilinguals' perceptions of feeling different when switching languages. *Journal of Multilingual and Multicultural Development*, *34*(2), 107-120.

Fishman, J. A. (1972). The Relationship between micro-and macro-sociolinguistics in the study of who speaks what language to whom and when. In Pride, J. B. & Holmes, J. (eds). *Sociolinguistics.* pp. 15-32. Harmondsworth: Penguin Books Ltd.

Gal, S. (1978). Peasant men can't get wives: Language change and sex roles in a bilingual community. *Language in Society*, *7(1):* 1-16.

Gal, S. (1979). *Language Shift: Social Determinants of Linguistic Change in Bilingual Austria*. New York: Academic Press.

Gammaldi, F. (2016). *Motivation of Code Switching in blogs of Italian expats living in the Netherlands* (Master's thesis). Retrieved from https://studenttheses.universiteitleiden.nl/handle/1887/37539

García, O., & Wei, L. (2014). Language, bilingualism and education. In *Translanguaging: Language, bilingualism and education* (pp. 46-62). Palgrave Pivot, London.

Gardner-Chloros, P., & Weston, D. (2015). Code-switching and multilingualism in literature. *Language and Literature*, *24*(3), 182-193.

Grim, F. (2008). The topics and roles of the situational code-switching of an English-French bilingual. *Journal of French Language Studies*, *18*(2), 189.

Gumperz, J.J. (1976/1982) "The sociolinguistic significance of conversational code switching". In Gumperz, J.J. (1982). *Discourse Strategies*. Cambridge: Cambridge University Press.

Gumperz, J. J. (1982). *Discourse strategies*. Cambridge University Press.

Heller, M., & Pfaff, C. W. (2008). Code-switching. In *Kontaktlinguistik* (pp. 594-609). De Gruyter Mouton.

Herring, S. C. (2008). Language and the Internet. *The international encyclopedia of communication*.

Hoffmann, C. (1991). *An Introduction to Bilingualism*. London: Longman.

Kieswetter, A. (1997). *Code-switching amongst African High School Pupils.* University of Witwatersrand Occasional Papers in African Linguistics.

Kroll, J. F., Dussias, P. E., Bice, K., & Perrotti, L. (2015). Bilingualism, mind, and brain. *Annual Review of Linguistics*, *1*(1), 377-394.

Ladegaard, H. J. (2018). Codeswitching and emotional alignment: Talking about abuse in domestic migrant-worker returnee narratives. *Language in Society*, *47*(5), 693-714.

Mabule, D. R. (2015). What is this? Is it code switching, code mixing or language alternating?. *Journal of Educational and Social Research*, *5*(1), 339.

Myers-Scotton, C. & Ury, W. (1977). Bilingual Strategies: The Social Functions Of Codeswitching. *International Journal of the Sociology of Language*, 5-20.

Myers-Scotton, C. (1983). The negotiation of identities in conversation : A theory of markedness and code choice. *International Journal of the Sociology of Language, 44*. 115-136

Myers-Scotton, C. (1993a). *Social motivations for codeswitching: Evidence from Africa.* Oxford University Press.

Myers-Scotton, C. (1993b) (1997). *Duelling languages: Grammatical structure in codeswitching*. Oxford: Clarendon Press for Oxford University Press.

Myers-Scotton, C. (Ed.). (1998). *Codes and consequences: Choosing linguistic varieties*. Oxford University Press.

Nilep, C. (2006). “Code switching” in sociocultural linguistics. *Colorado research in linguistics*, *19*.

Pavlenko, A. (2002). Bilingualism and emotions. *Multilingua*, *21(1),* 45-78.

Pavlenko, A. (2005). *Emotions and multilingualism*. Cambridge: Cambridge University Press.

Poplack, S. (1980). Sometimes I'll start a sentence in Spanish y termino en Español: Toward a typology of code-switching. *Linguistics, 18*, pp. 581-618.

Sebba, M., Mahootian, S., & Jonsson, C. (Eds.). (2012). *Language mixing and code switching in writing: Approaches to mixed-language written discourse.* Routledge.

Simin, A. M., & Hamid, B. A. (1994). English in Malaysia: The Case of two myths. In *Proceedings of the International English Language Education Conference*.

Stell, G., & Yakpo, K. (2015). *Code-switching between structural and sociolinguistic perspectives*. De Gruyter.

Wardhaugh, R. (1998). *An introduction to Sociolinguistics*. Massachusetts: Blackwell Publishers Inc.

Wei, L. (1998). *"The 'Why' and 'How' Questions in the Analysis of Conversational Codeswitching".* In P. Auer (ed.). *Code-Switching in Conversation: Language, Interaction, and Identity*. London: Routledge. pp. 156–76.

Weinreich, U. (1953) (1968). *Languages in Contact*. The Hague: Mouton de Gruyter.

Weston, D., & Gardner-Chloros, P. (2015). Mind the gap: What code-switching in literature can teach us about code-switching. *Language and literature*, *24(3),* pp. 194-212.

Woolard, K. (2005). "Codeswitching ". In A. Duranti (ed.). *A Companion to Linguistic Anthropology*. Malden, Mass.: Blackwell. pp. 73–94.

# Appendix A. Table

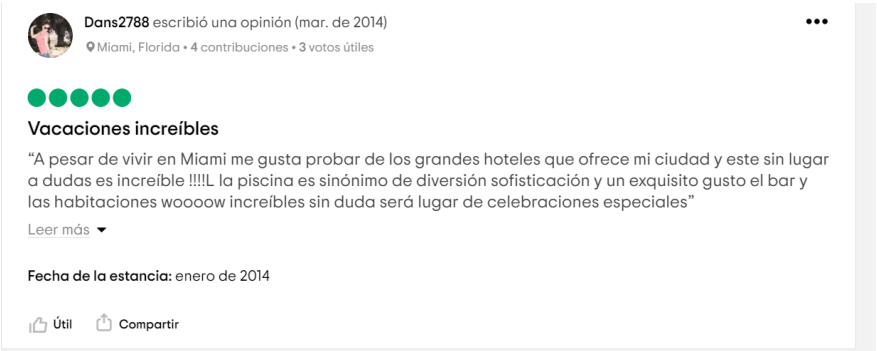
Table A.1. Chi-square distribution between final coding type of CS (1- inter-sentential, 2- intra-sentential, 3- tagging) and Review Rating (Negative, Neutral and Positive).

| | | | Negative | Neutral | Positive | Total |
|---|---|---|---|---|---|---|
| Final coding type of CS | 1 | Observed | 1a | 0a | 17a | 18 |
| | | % within final coding CS | 5.6% | 0.0% | 94.4% | 100.0% |
| | 2 | Observed | 8a | 10a | 55a | 73 |
| | | % within final coding CS | 11.0% | 13.7% | 75.3% | 100.0% |
| | 3 | Observed | 0a | 0a | 1[a] | 1 |
| | | % within final coding CS | 0.0% | 0.0% | 100.0% | 100.0% |
| | Total | Observed | 9 | 10 | 73 | 92 |
| | | % within final coding CS | 9.8% | 10.9% | 79.3% | 100.0% |

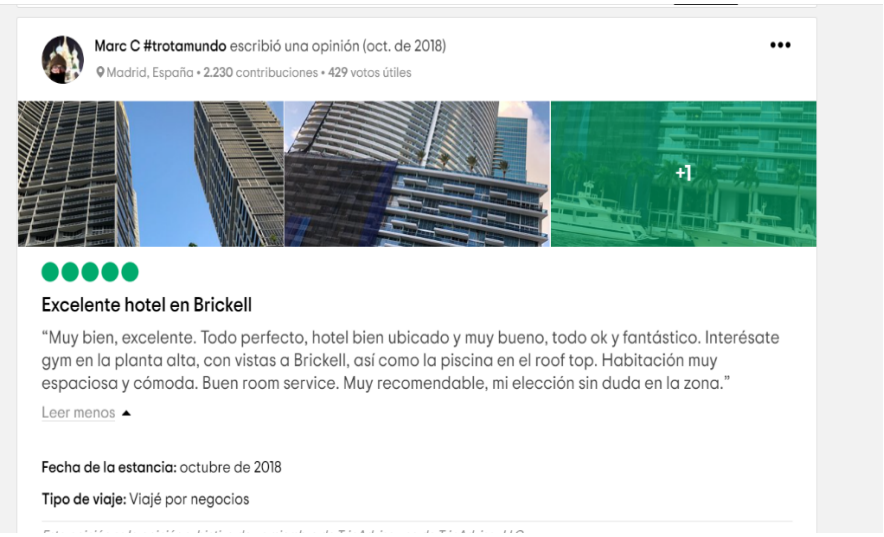
# Appendix B. Examples of Corpus reviews

Review 72:


Review 86:




Review 88:



# Appendix C. Categories of CS motivations

Category interpretation and explanation:

1. **Least effort principle** → The choice of the writer to spend as little effort as possible in their speech or text production (Barasa, 2010)
2. **Mode limitation** → The choice to encourage brevity, shorter words and/or shorter sentences (Barasa, 2010) .
    - For instance: Barbecue (English) and barbacoa (Spanish) are quite similar. However, there exists a short abbreviation in English, namely: BBQ. The Spanish language does not have an abbreviation like this, as a consequence, a CS to BBQ could be made to limit modes.
3. **Accuracy** → Some words simply do not exist in other languages or are hard to translate without losing ‘’meaning’’ (Barasa, 2010).
    - e.g. the Dutch word ‘gezellig’, does not have a direct equivalent in English and other languages
4. **Emphasis on message** → To highlight a concept/message or to catch attention (Myers-Scotton, 1993a)
5. **Communicating Irony** → Use of words to say the opposite of what you intend to say, often humorous (Myers-Scotton, 1993a)
6. **Distancing function** → Switching code to speak about topics that normally would be found too disturbing or upsetting in first language or to create distance between different ‘groups’ (Myers-Scotton, 1993a; Bond & Lai, 1986)
7. **Identity exposure** → Choice of language can have an impact on someone’s identity status and in the same way it can be a tool to communicate a certain (peer) identity (Barasa, 2010; Gammaldi, 2016)
8. **Exhibitionism** → Speaker/writer wanting to expose their linguistic skills (e.g. bilingualism) , class or identity in a way to show off to others (Gammaldi, 2016)
9. **Anglicism** → The use of easily understandable and common expressions or words that are borrowed from the English language, such as ‘wow’ (Gammaldi, 2016)

10. **Other** → For all the code-switches that you can not classify under the categories above.