5-2020

Effects of cognates on codeswitches

Ellie Decker
edecker2@vols.utk.edu

Follow this and additional works at: https://trace.tennessee.edu/utk_chanhonoproj

Part of the Spanish Linguistics Commons

Recommended Citation
https://trace.tennessee.edu/utk_chanhonoproj/2389

This Dissertation/Thesis is brought to you for free and open access by the Supervised Undergraduate Student Research and Creative Work at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Chancellor's Honors Program Projects by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.
Effects of cognates on codeswitches

Ellie Decker

University of Tennessee, Knoxville

Advisor: Dr. Harriet Bowden
Abstract

Cognates have been shown to trigger or facilitate codeswitching in proficient bilingual speakers in many studies. However, the proposed reason for this facilitation has not been examined with an electroencephalogram (EEG). Due to coactivation of languages, cognates are processed easier than noncognates. Previous studies have focused on single word codeswitches, effects of language dominance on codeswitching costs, and the relationship between cognates and codeswitches. However, the proposed study examines how initiating an alternational switch with a cognate affects its processing and the processing of subsequent words using EEG techniques. The brain’s electrical activity will be measured by the EEG as participants read codeswitched sentences. The examined sentences will contain a codeswitch from English to Spanish that is initiated either with a cognate or a noncognate and will be analyzed across multiple words. I hypothesize that codeswitches initiated with a cognate will require less processing of the codeswitched words (the initial and following words) compared to those initiated with a noncognate.
Effects of Cognates on Codeswitches

Codeswitching is the alternating from one language to another and is a common occurrence between two highly proficient bilingual speakers. Cultures of codeswitching commonly occur around borders of countries with different languages and between people of shared linguistic backgrounds. While codeswitching, speakers use correct grammar in both languages (Kroll et al., 2014). A speaker may either intentionally or unintentionally codeswitch depending on the type of codeswitch or reason for codeswitching (Broersma & De Bot, 2006). One example is a tag or lexical switch in which one word in the sentence is of a different language from the rest of the sentence. This is uncommon in natural codeswitching and more likely to be an intentional switch (Broersma & De Bot, 2006; Litcofsky & Van Hell, 2017). However, an alternational switch is when a sentence is a hybrid structure of two languages, containing one part in one language and another part in a different language. This type of codeswitch is observed more naturally and is more commonly triggered by cognates than a lexical switch (Broersma & De Bot, 2006).

Although switching languages usually results in greater processing requirements than remaining in the same language, many highly proficient bilingual speakers habitually codeswitch (Van Hell et al., 2015; Litcofsky & Van Hell, 2017). Most bilingual speakers have a dominant language to varying degrees (Litcofsky & Van Hell, 2017). A non-balanced bilingual is someone who uses one language more dominantly and effectively while a balanced bilingual is someone who does not have a dominant language.

Cognates are another element of bilingual language processing. A cognate is a word that shares a similar meaning, and pronunciation or spelling with a word in a different language. For example, tourist in English has the same meaning and similar pronunciation to turista in Spanish.
The special relation that cognates have between languages permit them to be processed differently. Cognates have been found to be processed quicker compared to non-cognates across multiple studies (Van Hell & Tanner, 2012). The processing of cognates has been found to be quicker in both the first and second language even in bilingual speakers who do not speak the second language fluently (Van Hell & Tanner, 2012). Many studies support that cognates result in a coactivation of both languages (Kroll et al., 2014). Additionally, cognates are considered to be trigger words for Codeswitching (Broersma & De Bot, 2006). For these reasons, cognates are of great interest when studying bilingual language processing.

Although cognates have been observed to facilitate codeswitching, more research may further explain the mechanism. The following study aims to illuminate the effect that initiating a codeswitch with a cognate may have on the processing of the codeswitch across multiple words. The hypothesis is that greater activation of the non-target language via cognates permits easier codeswitching since it may lower the language suppression of the language used for the subsequent words.

**Review of Literature**

**Bilingual Coactivation and Inhibition**

Many of the following effects are explained through the coactivation of languages. Coactivation is the simultaneous activation, or processing, of both languages in the brain. Both languages are understood to be activated at all times (Kroll et al., 2014). Although both languages are activated, inhibition of the non-target language occurs to allow for more attention toward or activation of the target language (Marian et al., 2017). The Inhibitory Control model assumes that people are able to use one language by suppressing the other (Van Hell et al., 2015). For example, in early second language learners the grammar and phonology of the first language
is at the forefront of their mind (Macizo, 2015). Macizo found that greater suppression is necessary when the speaker is more proficient in the non-target language than the target language in naming tasks (2015). With further proficiency, bilingual speakers are able to decrease or suppress the activation of the first language better when using their second language (Van Hell et al., 2015). The suppression of the non-target language is beneficial because it permits more attention toward the target language and less distraction from the non-target language. When comparing phonological competition within a language to between languages, more cortical resources are used to inhibit the non-target language (Marian et al., 2017). This further explains the additional costs of codeswitching because it requires a change to the previously suppressed language. Meanwhile, the previously selected language must then become inhibited.

The suppression of the non-target language results in greater working memory due to the greater use of executive control (Blom et al., 2014). In the Litcofsky and Van Hell study, more balanced bilinguals (i.e. similar proficiencies in both languages) performed better in the flanker task, a response inhibition task compared to more non-balanced bilinguals (2017). Bilingual children have shown greater ability in many working memory tasks compared to monolingual children (Blom et al., 2014).

Costs of Codeswitching

Many studies have found costs for codeswitching (Van der Meij, Cuetos, Carreira, & Barber, 2011; Litcofsky & Van Hell, 2017; Proverbio et al., 2004). These costs are due to processing extra information in order to understand a sentence that combines two languages (Van Der Meij et al., 2011). Specifically, even in fluent bilinguals that are habitual codeswitchers, there is greater cost when switching from the stronger language into the weaker
language than from the weaker language into the stronger language (Litcofsky & Van Hell, 2017).

The Litcofsky and Van Hell study concluded that this cost differential is due to the ability to process one’s dominant language easier as supported by the Inhibitory Control model (2017). This study was the first ERP study to utilize a switch involving multiple words to analyze the switch rather than one inserted word. Thus, it explained some broadened effects of codeswitching as used in more natural speech. The results reflected the greater processing when switching into the weaker language with greater amplitudes of the Late Positive Component (LPC) associated with sentence-level restructuring.

Another concern with codeswitching is how different types of codeswitches affect costs. A N400 effect appears depending on the type of codeswitch. Studies with a lexical codeswitch, or an insertion codeswitch which contains one codeswitched word, have greater amplitudes of N400s (Litcofsky & Van Hell, 2017). However, there were no N400 effects in the switches of the Litcofsky and Van Hell study which utilized alternational codeswitches (2017). In a reading aloud task, inter-sentential switching did not incur a cost whereas a lexical switch did (Gullifer et al. 2013). Inter-sentential and intra-sentential (which includes alternational codeswitches) codeswitches are more common and natural for bilingual speakers (Van Hell et al., 2015). This may mean that the uncommon occurrence of lexical switches will result in greater costs observed in the N400 which is associated with semantic and lexical processing (Van Hell et al., 2015). The results of these studies then indicate that alternational codeswitches may not result in additional lexical processing compared to non-switched sentences.

Some studies have found that there are conditions that prevent switching direction from having an effect on switching cost. Balanced bilinguals have been able to switch with equivalent
costs in either direction (Van Hell et al., 2015). While non-balanced bilinguals typically have greater costs switching from their stronger to weaker language, this cost differential diminishes when the non-balanced bilinguals are given a pre-switch cue (Van Hell et al., 2015). This suggests that with priming, non-balanced bilinguals may anticipate and be able to decrease language suppression of the non-selected language to compensate.

**Cognate Facilitation Effect**

The ability to process cognates quicker and more efficiently is largely due to the cognate facilitation effect (CFE). The cognate facilitation effect is the use of the shared or overlapping representation of a cognate between languages to facilitate its processing. Children as young as kindergarten have been shown to use the cognate facilitation effect (Rosselli et al., 2014). Among fluent bilinguals, there is greater use of the CFE in the weaker language as measured by reading time or response time for a cognate versus noncognate words. Non-balanced bilinguals use greater CFE than balanced bilinguals (Rosselli et al., 2014). Essentially, since cognates allow for easier processing, a non-balanced bilingual would rely on the CFE to process cognates easier in their weaker language.

For individuals learning a second language (L2), the Cognate Facilitation Effect may be a tool to reflect proficiency. Individuals learn cognates in their L2 faster than non-cognates. This observation is explained under the assumption that there is a transfer of vocabulary from one’s first language (L1) to one’s L2 (Rosselli et al., 2014). In the “Foreign language comprehension achievement: Insights from the cognate facilitation effect” study, the researchers used low-proficient and high-proficient L1 Spanish L2 English students to analyze the effect of cognate facilitation on learning (Casaponsa et al., 2015). It found that students with greater CFE during the initial stages of learning would develop greater proficiency than those students who did not
implement the CFE. However, the intermediate level students who over-relied on the CFE had lower proficiency (Casaponsa et al., 2015). This over-reliance represents a lack of L2 lexical development. These findings show that as language proficiency grows stronger, reliance on the CFE should decrease (Casaponsa et al., 2015). This is also reflected in the aforementioned study in which non-balanced bilinguals used more CFE than balanced bilinguals (Rosselli et al., 2014). Moreover, the CFE may hinder lexical abilities in non-balanced bilinguals. Midgley et al. demonstrated this with an ERP study that asked low proficiency second language learners to look at words in their L2 (2011). Cognates resulted in decreased amplitudes for the N400s compared to noncognates, meaning cognates required less lexical processing (Midgley et al., 2011). This indicated that participants inhibited the L2 more than the L1 (Midgley et al., 2011).

Additionally, bilinguals also perceive cognates differently than noncognates. Bilinguals perceive a higher frequency of a cognate word compared to monolingual speakers (Sherkina-Lieber, 2004). To bilinguals, cognates seem more frequent than noncognates even when frequency of noncognates and cognates are controlled (Sherkina-Lieber, 2004). This effect is called the Cognate Frequency Effect. The explanation for this effect is that since a cognate word has an overlapping or shared representation in the brain, it appears to be used more than noncognates which have their own representation separate in both languages (Sherkina-Lieber, 2004). Thus, the frequency is perceived greater. This is demonstrated by the “The cognate facilitation effect in bilingual speech processing: The case of Russian–English bilingualism” which concludes that the bilingual lexicon shares the representations of cognates (Sherkina-Lieber, 2004). This supports the coactivation model which argues that cognates activate both languages (Casaponsa et al., 2015; Kroll et al. 2014). Sherkina-Lieber also found that the higher frequency of a cognate in the L1 results in greater Cognate Frequency Effect in the L2 (2004).
This hints that the frequency effect may also be explained through the cognate facilitation effect and language inhibition. The dominant language is routinely less inhibited than the non-dominant language (Marian et al., 2017). Because words from dominant languages are less inhibited than words in the non-dominant language, considering cognates from the dominant language will be perceived in a greater frequency compared to considering cognates in the L2.

**Trigger Hypothesis**

A link between trigger words and codeswitches were observed in early studies in bilingual language production by Michael Cline (Broersma & De Bot, 2006). Trigger words are cognates, proper nouns, and borrowed words. While the original trigger hypothesis presumed a causal relationship between trigger words and codeswitches, Clyne later refined the hypothesis to characterize the trigger words as lexical facilitators of codeswitches rather than direct causes (Broersma & De Bot, 2006). While the presence of cognates make codeswitches more likely to occur, cognates are unable to predict the occurrence of a codeswitch (Broersma & De Bot, 2006). There is a correlational relationship between codeswitches and cognates.

In sentences containing trigger words and codeswitches, there is a relationship between the location of the trigger word to the location of the codeswitch. Codeswitches are most likely to occur directly after or directly before the trigger word (Broersma & De Bot, 2006). One explanation of this is the activation threshold hypothesis which determines that the lemmas of cognates activate the non-selected language (Broersma & De Bot, 2006). In other words, the order of lemma selection, which does not always follow the order of the sentence, determines the location of the switch (Broersma & De Bot, 2006). Thus, a codeswitch may occur before a trigger word in speech since the lemma of the trigger word may have been selected prior to when the lemma of the codeswitch word was selected. The non-selected language will be most
activated immediately after the selection of the lemma of the trigger word (Broersma & De Bot, 2006). Broersma and De Bot adjusted the trigger hypothesis to reflect this relationship in that cognates simply caused the previously suppressed non-selected language to be more activated (2006). The closer the activation of the non-selected language is to the activation of the selected language, the more likely a codeswitch is to occur.

**Motivation**

The goal of this research study is to develop a greater understanding of the causes for the cognate facilitation of codeswitches as explored in the trigger hypothesis and how a cognate codeswitch is processed in an alternational switch.

Previous studies have not investigated a cognate at the location of an alternational codeswitch with ERP data. Although studies have shown that cognates do facilitate a codeswitches and make codeswitches easier to process, there has not been an ERP study to determine its possible effects on subsequent words. Given the decreased inhibition of non-selected language when processing a cognate, the use of a cognate to initiate an alternational codeswitch may aid in providing more support to the trigger theory as adjusted by Broersma and De Bot (2006). When considering that one issue in analyzing codeswitches is that the lemma order cannot be predicted during production studies, the visual stimuli in this study will be presented one word at a time which forces the lemma of the cognate to be processed before considering the following words. Although this study does not follow the specifications of a stereotypical codeswitch in which the codeswitch falls after the cognate, the codeswitch instead falls on the cognate which serves another purpose in elucidating how a cognate is processed in an alternational codeswitch. The proposed study may indicate that cognates make processing
codeswitches easier by decreasing cross-language inhibition for the words following it with ERP data.

Although Van Der Meij et al. (2011), Moreno et al. (2002), and Proverbio et al. (2004) found N400s for lexical codeswitches, Litcofsky and Van Hell’s results lacked any effects on N400s (2017). Litcofsky and Van Hell determined this to be due to the difference that alternational switches (that are not initiated with a cognate) do not produce N400s where lexical switches do (2017). However, there may be a difference between the switches in the proposed study considering that cognates have decreased N400 amplitudes compared to noncognates (Midgley et al., 2011). Given the multiple factors involved, analyzing the time frame of a N400 will be necessary to further evidence that cognates result in decreased costs of lexical processing. However, Litcofsky and Van Hell’s study demonstrated that alternational codeswitches would result in LPCs which reflect sentence-level restructuring (2017). For this reason, analyzing the LPC will be necessary to demonstrate that through removing the suppression of the non-selected language, cognates are able to decrease the sentence-level restructuring in alternational codeswitches. Another possible observed effect may be the Late Anterior Negativity (LAN effect) which has been observed in some studies, including Moreno et al. (2002; Litcofsky & Van Hell, 2017).

Methods

Participants

Participants (n=6-10) will be heritage speakers of Spanish, college aged, and have an age of acquisition of English around 6. Participants will be right-handed, have normal or corrected vision and hearing, and have no neurological or psychiatric disorders. Because the EEG portion of the study utilizes visual stimuli, it is important for participants to be at least moderately
familiar with reading in Spanish. For this reason, eligible participants will be minors or majors in Spanish with 3 or more upper division Spanish courses (300+ level) completed. Due to the existence of shared cognates between multiple Romance Languages, participants will not be eligible if they have had learning experiences with other Romance Languages.

**Materials**

*Language Background Materials and Proficiency Tasks*

In order to determine language dominance, many measures will be taken to assess proficiency and usage in both English and Spanish. The phone screening will solidify necessary language background information such as age of acquisition, time spent in countries where English is not the dominant language, and Spanish courses taken. A language background questionnaire will be used to understand the breakdown of Spanish and English use, based on the *Language and Social Background Questionnaire* (LSBQ) (Anderson et al., 2018). The version created for this study is shorter and excludes information that would be gathered during the phone screening.

A picture naming task will be completed as another measure of proficiency. Black and white pictures from the International Picture Naming Project (IPNP) database were used to create a task with a Spanish and English version (Bates et al., 2000). Each version consists of 30 different pictures presented in order from most to least common objects. This was determined based on the word’s frequency and correct answer percentage from participants of the IPNP (Bates et al., 2000). The test begins with the object with the highest correct answer percentage and word frequency before decreasing in both as the objects later in the task are less common. The word frequency and correct naming percentages were balanced between the Spanish and English version to be significantly similar.
A self-paced reading task was constructed from sentences used in the Litcofsky and Van Hell study that are unutilized in the below EEG stimuli (2017). Four versions of this task were created from 24 total sentences in which a participant will read 12 sentences in English and 12 sentences in Spanish in which no sentences are repeated across language. Two questions per version will be asked after completion to ensure comprehension and attentiveness.

To prevent effects of priming, none of the above tasks, including pictures or words within the instructions, tasks, or questions, contain critical words for the EEG stimuli. However, this does not include prepositions and articles which may be in critical positions.

*Cognitive Tasks*

There are two cognitive tasks. A Flanker task from the online software PsyToolKit will be used to measure response inhibition which plays a role in executive function, important in both codeswitching and language suppression (Stoet, G., 2010; Stoet, G., 2017). An operation-span task will measure working memory.

*EEG Stimuli*

EEG stimuli was taken from or based on the 2017 Litcofsky and Van Hell study. The sentences used in this study, however, only switch in one direction from English to Spanish in which every sentence has a version with a noncognate and a cognate as the first codeswitched word. An example of these two versions as well as a filler sentence is shown in Figure 1.

<table>
<thead>
<tr>
<th>Cognate Switch</th>
<th>All summer long, Chris collected spiders and other insectos para examinar en su microscopio nuevo.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3</td>
</tr>
<tr>
<td>Noncognate Switch</td>
<td>All summer long, Chris collected spiders and other muestras para examinar en su microscopio nuevo.</td>
</tr>
<tr>
<td></td>
<td>1 2 3</td>
</tr>
<tr>
<td>Filler Sentence</td>
<td>Yesterday we agreed that the striped pillows match the rug in the living room.</td>
</tr>
</tbody>
</table>

*Figure 1* Three types of sentences will be used for the EEG stimuli. The two conditions are the cognate switch and the noncognate switch. The 1st, 2nd, and 3rd codeswitched words are to be analyzed and are labeled 1, 2, and 3. An example of a filler sentence is at the bottom.
There are 60 base sentences (120 sentences in total) and 60 filler sentences. As in the Litcofsky and Van Hell study, all sentences are between 10-22 words long with the codeswitch varying in position from word 7-16 (2017). The codeswitched sentences were checked by a native Spanish speaker who works as a Spanish professor, one heritage speaker who considered herself to be a habitual codeswitcher, and one heritage speaker who considered himself to be a non-habitual codeswitcher. After corrections, all three determined the grammar to be correct and believed that the codeswitches fell in natural locations. All filler sentences are in English and lack codeswitches. The change in location of the codeswitch as well as the filler sentences function to prevent predictability of the codeswitches. All 60 cognates and 60 noncognates were balanced to be similar in word length (number of letters), lemma frequency, and surface frequency. Neither cognate or noncognate words contain any diacritic marks such as accents or tildes, and cognates do not share exact spelling in order to differentiate it from the English. Word frequency counts were obtained from the Corpus del Español: NOW database (Davies, 2019).

Sentences were pseudorandomized to prevent more than 4 codeswitched or 4 filler sentences from appearing in a row. For one of every eight sentences, a true or false comprehension question will appear on screen. These comprehension questions were placed in a pseudorandomized order so that the question appears at a different location in every grouping of 8 sentences. There are an equal number of comprehension questions for filler sentences as for codeswitched sentences in order to prevent extra focus or attention on the codeswitched sentences as well as distract from the purpose of the study. A rest screen will occur between the presentation of each sentence to allow participants to blink and prepare for the next sentence.

Procedure
The study will consist of one session. Participants will complete an informed consent form and a language background questionnaire. The operation-span task and the flanker task will then be completed on a computer followed by both the Spanish and English versions of the picture naming task. For the picture naming task, participants will be asked to name the picture that appears on the screen into a microphone attached to a serial response box to measure response time and accuracy. Participants then complete the self-paced reading task. Next, Participants begin the EEG portion of the study where they read sentences silently and answer comprehension questions. The sentences are presented one at a time in the center of the screen, each word appearing for 200 ms, followed by 500 ms before the presentation of the following word. The final word of each sentence appears on the screen for 500 ms. The sentences are organized in four blocks of 40 sentences each, allowing for a break between each block.

A 64-electrode WaveGuard cap will record electrical brain activity with a sampling rate of 512 Hz. Two horizontal and two vertical electrodes will observe horizontal and vertical eye movements. Exclusion of artifacts in trials (e.g. blinking) will prevent disturbances of data. Data from participants who score lower than 70% on the comprehension questions will be excluded as well.

**Data Analysis**

In order to determine language dominance, all proficiency and language use measures (Language background information, self-paced readings, and the picture naming task) will be scored. Participants who received higher scores in one language were assumed to be dominant in that language. One-way ANOVAs will be used to analyze the mean and standard deviations of these results. An ANOVA will also be utilized to determine the results of both cognitive tasks.

MatLab will be used to examine the EEG data. The average waveforms from the time frame of interest (200-1000 ms) will be used to compare the independent variable of the status of
the initial word’s effects in pairs: the first codeswitched word (the cognate compared to the noncognate), the second codeswitched word (comparing the word following the cognate and noncognate), and the third codeswitched word (comparing the second subsequent words to the cognate and noncognate). The first, second, and third codeswitched words are labeled 1, 2, and 3 of the given condition on Figure 1. Given the results from previous studies, the primary components of interest are the N400 and LPC. To watch for a potential LAN, scalp maps will be investigated between 250 ms and 550 ms. Because working memory, response inhibition, and language dominance have an effect on codeswitching, ANOVAs will be performed to analyze any independent correlations or effects on the critical components.
References


