Neuronal activity to environmental sounds when presented together with semantically related words

An MMN study on monolingual and bilingual processing of homophones

Yeliz Afyonoglu Kirbas
Neuronal activity to environmental sounds when presented together with semantically related words

An MMN study on monolingual and bilingual processing of homophones

Yeliz Afyonoglu Kirbas

Abstract

Neuronal activity of monolingual and bilinguals to environmental sounds and words that are semantically related to them were studied using the mismatch negativity (MMN) component of event-related potentials. MMN was expected to reflect the language selection process in bilinguals on the bases of semantics and phonology. In this regard, as lexical stimuli, interlingual homophones ‘car’ and ‘kar’ (snow) were presented together with semantically related environmental sounds in a passive auditory oddball paradigm. The lexical stimuli were recorded by a native English speaker. Three Turkish-English late bilinguals and one native English speaker participated in the study. An early MMN was elicited in both groups with a distribution over the fronto-central and central areas across the scalp with highest peak amplitude at ~2.5 with a 113 ms latency. This response indicates that participants of the study were sensitive to the acoustic changes in two different types of stimuli. The further investigation of the interplay between environmental sounds and semantics displayed no conclusive result due to lack of data. Finally, the brain responses gathered from the bilinguals were not enough to draw a conclusion. Whether the bilinguals were sensitive to the sub-phonemic cues in the presented auditory lexical stimuli or not were inconclusive.

Keywords
Environmental sounds, interlingual homophones, mismatch negativity, event-related potentials.
Neuronal aktivitet av miljöljud när de presenteras tillsammans med semantiskt relaterade ord

En MMN studie i en- och tvåspråkig bearbetning av homofoner

Yeliz Afyonoglu Kirbas

Sammanfattning


Nyckelord

Miljöljud, interlingual-homofoner, mismatch negativity, event-relaterade potentialer.
## Contents

1 Introduction ................................................................................. 1

2 Background ............................................................................... 2

  2.1 Bilingualism ......................................................................... 2
  2.2 Interlingual homophones ...................................................... 2
  2.3 Semantic memory ............................................................... 3
      2.3.1 Priming ......................................................................... 4
  2.4 Environmental sounds (ES) .................................................... 4
  2.5 Languages of interest ......................................................... 5
  2.6 Event-Related Potentials (ERPs) ........................................... 6
  2.7 The Mismatch Negativity (MMN) ......................................... 6
      2.7.1 Previous MMN studies ................................................. 7
  2.8 Other ERP components of interest ....................................... 8
      2.8.1 N1 - P2 complex .......................................................... 8
      2.8.2 N200 complex ............................................................ 8

3 Purpose and research questions ................................................. 9

  3.1 Research questions and hypotheses ..................................... 9

4 Method and material ............................................................... 10

  4.1 Participants .......................................................................... 10
  4.2 Stimuli ................................................................................ 10
  4.3 Setting up the EEG ............................................................. 11
  4.4 Procedure ........................................................................... 11
  4.5 EEG Data Analysis ............................................................. 12

5 Results .................................................................................... 13

  5.1 ERP data for the L1 English speaker .................................... 13
  5.2 ERP data for the Turkish-English bilinguals ....................... 14
  5.3 Topographical distribution ............................................... 15

6 Discussion .............................................................................. 16

  6.1 Result Discussion ............................................................... 16
  6.2 Method Discussion ............................................................ 18
  6.3 Further Research ............................................................... 19

7 Conclusion .............................................................................. 20

References .................................................................................. 21
1 Introduction

The interplay between environmental sounds and words that are semantically related to them are intriguing. Environmental sounds, although being non-linguistic elements, provide information about one’s surroundings by referring to real-life objects. When hearing an environmental sound such as 'meow', the fact that there is possibly a cat around is not difficult to recognize for the human brain due to the natural and causal relationship between the two stimuli. That is, environmental sounds can function as non-arbitrary references for words (Van Petten & Rheinfelder, 1995, p.485). However, this conceptual relationship can be thought as language-dependent, since hearing the word ‘cat’ would not lead to the same representations for a listener who does not speak English. Given that there is evidence for the effects of non-linguistic information on the interpretation of the linguistic element and that there is co-activation for lexical representations in bilinguals, a study design involving both bilingual and monolingual participants, in which the stimuli are presented both as linguistic and non-linguistic elements in a comparable manner, may lead to a better understanding of how the relationship between these stimuli influence the brain processes. In this regard, in the present study, both mono- and bilinguals were investigated when hearing environmental sounds and phonologically similar words that are semantically related to them. That is participants in this study were presented with the English word ‘car’ and the Turkish word ‘kar’ meaning ‘snow’ as auditory lexical stimuli and environmental sounds that are semantically related to them randomly while watching a silent documentary.

Whether it is a speech sound or a non-linguistic stimulus such as environmental sound, language processing begins with acoustic-phonological analysis, which occurs in a couple of milliseconds after stimuli onset (Friederici, 2017, p.20). To be able to capture such a fast process, the time-sensitive neurophysiological technique of electroencephalography (EEG) was used, which gives information about when things happen in the brain in millisecond precision. So, with this method at hand, it became possible to focus on a component known to be related to early language processing, namely the Mismatch Negativity (MMN). Due to MMN’s intrinsic properties of being an index of brain mechanisms of speech perception and understanding (Näätänen, 2007, p.135) and furthermore, MMN’s being a robust indicator of change detection in auditory stimuli, firstly, whether an MMN response be elicited or not was asked in a design where the two stimuli are distinct (i.e., speech sound in contrast to environment sound). Secondly, since MMN is known to reflect language-specific memory traces (Näätänen, 2007, p.135), it was also questioned how those stored language representations would influence brain responses of monolinguals and bilinguals when the two stimuli are semantically related or unrelated environmental sounds and words. Finally, because the presented auditory Turkish word stimuli involve sub-phonemic cues it was asked whether bilingual participants would access their L2 when the auditory word stimulus is presented in their L1 or not.
2 Background

2.1 Bilingualism

Learning a new language as a child, a teenager or an adult is argued to affect the brain functions positively (Maher, 2017, p.66) and having a language competence in two languages can be defined as bilingualism (Maher, 2017, p.62). However, the so-called ‘ideal equilingualism’ or balanced bilingualism is rare, in which the bilingual person has balanced competence in both languages, and native-like controls for a bilingual in both languages mostly vary (e.g., listening over reading, writing over speaking, speaking over writing) (Maher, 2017, p.62-3).

One of the interesting research questions in bilingualism is whether two languages of a bilingual activate in parallel during lexical access. One of the views is the so-called language non-selective access. According to this view, a bilingual that speaks language X and language Y would activate word W in both languages, that is there is evidence for co-activation for lexical representations in bilinguals (Dijkstra, Grainger, & van Heuven, 1999, p.510; Schulp, Dijkstra, Schriefers, & Hasper, 2003, p.1172). Furthermore, the language non-selective access view assumes that, when a stimulus is presented in only one of the languages, both languages would be activated in parallel in the bilingual brain (de Bruijn, Dijkstra, Chwilla, & Schriefers, 2001, p.155) in contrast to the language selective-access view, which holds that only the relevant language is accessed (Schulp et al., 2003, p.1155).

The first step to online language processing begins with hearing. The auditory system in healthy hearing individuals responds to any type of stimuli from a simple tone to a loud noise such as an engine running or to a speech sound (Urbach & Kutas, 2018, p.7). Whether it is a speech sound or not, the auditory input first enters the peripheral auditory systems and the processing for the given stimuli begins in the brain (Zora, 2016, p.31). The stock of words a person knows is called the mental lexicon and the retrieval of those words’ representations from the long term memory is known as spoken word recognition (Vitevitch et al., 2018, p.1). So, if the presented stimulus is lexical, the word recognition occurs by retrieving the lexical representations from the so-called mental lexicon depending on the speech sound’s structure such as segmental information (e.g., vowel, consonant) as well as suprasegmental information (e.g., stress and tone) to match the ones that are stored in the long term memory (Vitevitch, Siew, & Castro, 2018, p.1; Zora, 2016, p.19).

Lexical retrieval in bilinguals, however, is argued to be weaker (Bialystok, Craik, & Luk, 2008, p.535). Previous studies show evidence for longer response times in bilinguals, regardless of their dominant language, in comparison to monolinguals (Bialystok et al., 2008, p.523; Bialystok, Craik, & Luk, 2012, p.2). However, these effects were also discussed to be in relation to the possible smaller size of the vocabulary in the second language than the native language (Bialystok et al., 2008, p.535). In a study conducted with German children who emigrated to Sweden, bilingual children were slower in object, word and number naming tasks than their monolingual peers in school (Mägiste, 1992, p.363). However, when the naming task was conducted with cognates, bilinguals preceded the monolinguals (Libben et al., 2017, p.30).

The executive control in bilinguals on the other hand, is known to be better than monolinguals, that is, they are known to have better cognitive resources for working memory, inhibition and attention switch (Bialystok, 2010, p.93). In a study conducted with Chinese-English bilinguals, Thierry & Wu (2007) investigated the inhibition control and found that the participants spontaneously accessed Chinese translations when reading or listening to English words. Accordingly, they argue that “[n]ative-language activation is an unconscious correlate of second-language comprehension” (p.12533).

2.2 Interlingual homophones

For the above mentioned language non-selective access view, most of the evidence comes from cross-linguistic studies, where the auditory or visual stimuli are words that mean the same such as cognates,
or words that share their orthography such as homographs or share their phonology like homophones (Novitskiy, Myachykov, & Shtryov, 2019, p.210). Words that sound almost the same are known as homophones (e.g., pair-pear) and when this phenomenon occurs between languages it is called ‘interlingual homophones’ (e.g., ‘boec-bouc’ /buk/ a Dutch-French homophone pair meaning book and goat respectively) (Schulpen et al., 2003, p.1155).

According to one of the bilingual auditory word recognition models, Bilingual Interactive Activation Model of Lexical Access BIMOLA, in the processing of ambiguous words such as homophones, both language representations will be activated and inhibition can be a result of an external effect such as being in an irrelevant environment (Schulpen et al. 2003, p.1171-2). In this model, being a non-selective language model, the heard word stimulus activates common phonological features of the bilinguals’ two languages, and these features then activate phonemes. After that, it is either the activation of subset of other phonemes that belong to the same language with the given input or an inhibitory effect can take place. Homophones have been interpreted differently over their outcome on whether they have an inhibitory or facilitative effect (Carrasco-Ortiz, Midgley, & Frenck-Mestre, 2012, p.532).

In their study, focusing on priming effects, Schulpen et al. (2003) investigated both offline and online processing of interlingual homophones in Dutch-English late bilinguals. They presented participants first the auditory and then the visual stimulus. These stimuli were either related (e.g., /li:s/ - lease) or unrelated prime-target pairs (e.g., /pig/ - lease). In addition to that, there was a third condition where participants were presented interlingual homophones in Dutch and English (e.g., /li:s/ - lies in Dutch) and (e.g., /li:s/ - lease in English). The stimuli were recorded with an English pronunciation by a Dutch-English bilingual who lived in England for many years (p.1158-1165). Schulpen et al. (2003) expected faster response times for the English homophone ‘/li:s/ - lease’ pair than the Dutch ‘/li:s/ - lies’ homophone pair due to the cues that was in the presented stimuli, that is the English pronunciation of the word (p.1159). Their results were in line with their expectations and they found that the auditory lexical access was non-selective in bilinguals. They also expected competition in recognition for the interlingual homophones, since homophones are harder to identify which leads to longer recognition times in comparison to monolingual controls (p.1170). However, the authors argued that homophones, unlike written words, carry more information so that the listener can distinguish the presented auditory stimuli with the help of sub-phonemic cues such as vowel length, VOT, and aspiration (p.1156). According to them, these cues, which are known to be language specific and acquired with experience (White, 2018, p.22), helped the bilinguals to resolve the process of choosing one interlingual homophone over the other (Schulpen et al. 2003, p.1170). Their results showed that, with the English pronounced words as stimulus, participants decided faster for the English homophone ‘lease’ than the Dutch homophone ‘lies’ after hearing the interlingual homophone /li:s/. Schulpen et al. (2003) concluded that their participants used the sub-phonemic cues to make decisions in regard to ambiguous stimuli like interlingual homophones to select the relevant language (p.1171).

### 2.3 Semantic memory

Memory is vital for language processing in regard to encoding, storing and retrieving information (Heredia & Altarriba, 2014, p.4). There are three basic memory systems: (i) sensory memory, in which the input first enters with and the information is stored very brief moment of time; (ii) short-term memory, which has a small capacity to store information for very near future; and (iii) long-term memory, in which the information can be stored for a long period of time (Heredia & Altarriba, 2014, p.5). Semantic memory is known to be a long-term memory system, in which meanings of words, facts such as a mathematical formula or general knowledge are stored (Heredia & Altarriba, 2014, p.133). Several models have been suggested to explain semantic memory such as experience-based and distributional-based semantic memory models. According to experience-based model, the word duck is represented with its most common features (e.g., quacks, has fins, swim in the water). Distributional model, on the other hand, holds the view that “[t]he co-occurrence of words is thought to form the basis of word associations” (Heredia & Altarriba, 2014, p.136).
As for the bilingual memory, it is argued that the concepts are shared as an underlying system of separate lexicons that belong to the bilingual’s languages (Heredia & Altarriba, 2014, p.2), which is known to be a dynamic system that depends on language use and language dominance (Heredia & Altarriba, 2014, p.29). Among the influential models of bilingual memory are Word Association Model and Concept Mediation Model. According to these models, word representations are stored independently of language at the lexical level and semantic representations at the conceptual level (Heredia & Altarriba, 2014, p.147-8). In Word Association Model, the L2 word meaning in the conceptual level is only accessed through the activation of L1 and in the Concept Mediation Model L1 and L2 are connected to conceptual level separately (Heredia & Altarriba, 2014, p.148). A combination of these two models is the so-called Revised Hierarchical Model, in which the connections between L1 and L2 can differ, that is L2 proficiency is argued to affect the lexical links between the two languages separately (Heredia & Altarriba, 2014, p.150).

### 2.3.1 Priming

Priming is known to be a non-conscious form of memory (Tulving & Schacter, 1990, p.301). It occurs by tapping in a person’s memories by saying a single word (Bosch & Leminen, 2018, p.462) and this way one provokes a target word or word sequences. However, it also differs individually due to the continuous experiences one has in relation to the word (Hoey, 2012, p. 8, 18). For example, when a listener hears the word ‘table’ recognition of the word ‘chair’ would be faster than the word ‘baby’, which is an unrelated word to the word ‘table’ for many people probably. Thus, priming is argued to ease memory’s tasks (Hoey, 2012, p.8).

Priming is also known to shape the language, and bilingual memory explanations come mostly from cross-linguistic priming studies (Heredia & Altarriba, 2014, p.28). Hoey (2012) claims that “everything we know about a word is a product of our encounters with it” (p.2). So, every experience builds up new associations to a word by confirming or by weakening (Hoey, 2012, p.12). Neurologically, associative learning is explained as the forming of a network due to the strengthening of mutual connections in two different brain areas that activate simultaneously (Shtyrov and Hauk, 2004, p.1089). It was Hebb (1949) who coined the well-known phrase ‘neurons that fire together wire together’. Hebb (1949) stated that “[a]ny two cells or systems of cells that are repeatedly active at the same time will tend to become ‘associated’ so that activity in one facilitates activity in the other” (p.70) which then generates mental representations for the related stimuli such as words and objects with the so-called cell assemblies (Lansner, 2009, p.178).

In a study conducted with Spanish-English late unbalanced bilinguals and English monolinguals, Hoshino & Thierry (2012) investigated semantic priming with interlingual homographs where the targets preceded by primes, which were either related to the English or Spanish meaning (p.1). Their results showed that monolinguals were affected by priming when primes were related to the English homograph, which is their L1. Bilinguals, on the other hand, showed no priming effects to either of the languages (p.5).

In the present study, the stimuli have a semantic relationship. They are the English noun ‘car’ and the environmental sound ‘car engine/motor’ (hereby engine) as well as the Turkish noun ‘kar’, which means “snow” and the environmental sound ‘rain’. In logical terms, word A is a meronym of word B if and only if the meaning of A names a part of B. So, the English concrete noun ‘car’ has a meronym relation with the word ‘engine’, and is expected to be primed when presented together. On the other hand, the Turkish noun ‘kar’ is a weather condition and can be associated with other weather conditions such as rain. Snow and rain have a taxonomic sisterhood relation, both being hyponyms of weather condition.

### 2.4 Environmental sounds (ES)

Environmental sounds (ES) are usually defined as sounds generated by real events such as bell ringing, water splashing, baby crying, cow mooing, cat hissing and many more which refer to a very small set of referents unlike the speech sounds that are arbitrarily assigned to their referents
Invent the Turkic language family members are, uvular, alveolar trill, tap, post alveolar and it is also considered as being voiced post-alveolar. According to the lexical database of English WordNet (2010), the concrete noun words. As a result, the authors claim that the conceptual relationships between spoken words and ES are related in meaning influence the processing of both words and sounds (p.504).

Cummings et al. (2006) conducted a similar study to investigate the processing of ES with nouns and verbs in an audio-visual, cross-modal sound–picture match/mismatch paradigm (p.94). First, they presented participants a visual stimulus, a picture on a screen and then whilst looking at the picture participants heard an auditory stimulus either matching the picture or not. After the sound offset, participants were asked to press a button to indicate whether the picture and the sound match or mismatch (p.103). They found the sound type and word-class interaction, in which the response times for noun stimuli were faster for matched auditory words and the ES. Furthermore, Cummings et al.’s(2006) data displayed earlier effects in brain responses in recognition of ES than words. As a conclusion, they claimed that the ES directly activate related semantic representations and their processing may be faster due to their internal acoustic variability (e.g., bass & treble ratio, strength) compared to the processing of lexical stimuli (p.100).

### 2.5 Languages of interest

In the present study, the investigated languages are English¹ and Turkish, and the investigation is limited only to the features that are relevant for this study. The presented lexical stimuli are interlingual homophones ‘car’ /kaːr/ from English and ‘kar’ /kar/ from Turkish. According to the lexical database of English WordNet (2010), the concrete noun ‘car’ is a vehicle; usually propelled by an internal combustion engine. The word ‘kar’ in Turkish is also a concrete noun meaning “snow”.

According to the World Atlas of Language Structures (WALS), West Germanic language of the Indo-European language family, English has 24 consonants in its inventory when excluding the sounds that are used only in borrowed words. Among the consonants, the liquid /r/ is argued to have the most phonetic variants among the other consonants of English (McCully, 2009, p.44). Some of its variants are, uvular, alveolar trill, tap, post alveolar and it is also considered as being voiced post-alveolar approximant. Moreover, it is the only retroflex in American English (Yavas, 2011, p.7).

The Turkic language family member Turkish is a typologically distinct language from English and has 29 letters in its alphabet. According to the University of California Los Angeles Phonological Segment Inventory Database (UPSID) (2019), the total amount of segments in Turkish are 33.

---

¹ In this paper the discussed English language is American English.
Being a phonemic language, Turkish has correspondence between letters and sounds higher than English, however, their pronunciation may vary depending on the vowel that precedes or follows them (Göksel & Kerslake, 2011, p.2). One relevant feature to the present study is plosives. For example, the voiceless velar plosive /k/ is aspirated as onset, in initial position of a word when followed by a vowel. Similar aspiration pattern is also common in American English (Yavas, 2011, p.58). Furthermore, the articulation of /k/ varies depending on the vowel which it precedes, so if it is followed by a back vowel as in the word [kar] “snow”, the articulation is affected, and the tongue moves backward because of the back vowel /a/ (Göksel & Kerslake, 2011, p.3). According to Yavas (2011), American English has a similar feature for velars (p.62).

One of the most problematic sounds is the consonant ‘r’. Göksel & Kerslake (2011) describes the production of the Turkish /r/ as: “[b]y touching the tip of the tongue on the medial part of the palate” (p.2) and Yavas (2011) argues that there is a significant difference can be found in liquids such as the alveolar tap /s/ in final position. According to him, even though the distributions are different they create mismatches (Yavas, 2011, p.191).

Vowel sounds are classified according to their height, backness and lip-rounding (Maddieson, 1984, p.123). Thus, the tongue is primary as well as the lips’ and the jaw moving are necessary for their articulation. The vowel in the English and Turkish words are unrounded back vowel [a]. In the presented stimuli, the Turkish word ‘kar’ was sounded as [v], however, this does not change the meaning of the word.

2.6 Event-Related Potentials (ERPs)

The non-invasive neurophysiological method to record electrical activity from the scalp is known as electroencephalography (EEG). The EEG recordings show neuronal activity related to language processing in a millisecond resolution. However to distinguish the investigated processes from the other ongoing activities in the brain that are unrelated to the investigated phenomena is difficult (Luck, 2014, p.4). Event-related potentials, on the other hand, make it possible to distinguish the sensory, cognitive and motor responses from the overall collected EEG data. ERPs are electrical potentials that are time-locked to specific events as Luck (2014) describes them (p.4). They originate mostly as post-synaptic potentials (PSP) which occur within a single neuron (Luck, 2014, p.12-3). These PSPs create a flow of tiny electrical current, a dipole, and ERPs can only be measured from the scalp when these single neurons sum up and travel together to the same destination as a response to a specific event (Luck, 2014, p.13).

Some of the ERP components are known as language related such as N400, which is sensitive to semantic violations (Kutas & Federmeier, 2000, p.463), and P600, which is sensitive to syntactic violations such as gender agreement (Urbach & Kutas, 2018, p.14). In the present study, the focus will be on the Mismatch Negativity (MMN) component of ERPs. Previous studies show evidence for MMN being sensitive to all linguistic processes (Shtyrov & Pulvermüller, 2007, p.179; Zora et al., 2015; Zora et al., 2016a; Zora et al., 2016b).

2.7 The Mismatch Negativity (MMN)

ERP investigations focusing on the brain mechanisms involved in auditory processing is widely used in cognitive and psycholinguistic studies. One of the robust ERP components in auditory processing is the so-called Mismatch Negativity (MMN). This widely accepted event-related potential, which is elicited by any discriminable auditory-change in stimulation, is also a well-known index of long-term memory traces in human brain (Shtyrov & Hauk, 2004, p.1089). MMN is mostly investigated with the so-called oddball paradigm, which was first introduced by (Squires, Squires & Hillyard,1975). It is an experimental design, in which a frequently repeated stimulus ‘standard’ is interrupted with a rarely repeated stimulus. These rare stimuli are called deviants/oddballs. The standard stimulus forms a representation, a memory trace for the heard sound for a few seconds, and this short-term memory trace is violated with the presentation of deviant stimulus resulting in a MMN response (Näätänen et
al., 2007, p.2545-6, 2548). MMN is acquired by the calculation of the difference between the two, that is by subtracting the standards from the deviants. The deviations can be in the basic acoustic features such as pitch, intensity, duration and frequency with respect to the frequently presented stimuli (Shtyrov & Pulvermüller, 2007, p.176).

MMN has a frontocentral scalp distribution (Näätänen et al., 2007, p.2549). It usually peaks at 150-250 ms from change onset that is sensitive to the magnitude of the change in the stimulus which affects the latency of the peak, resulting in an earlier threshold as early as 100 ms (Näätänen et al., 2007, p.2545).

The pre-attentive feature of MMN makes it widely applicable since there is no requirement of a task for it to be elicited. That is, MMN is elicited regardless of participant’s attention and considered as being an automatic component (Shtyrov & Pulvermüller, 2007, p.178). In fact, it has been argued that its amplitude gets larger with the attention directed away from the intended stimuli. To draw the attention away from the intended stimuli, simultaneous tasks are usually applied (e.g., watching a silent movie, reading a book) (Näätänen, 1984, p.286; Tamminen, Peltola, Kujala, & Näätänen, 2015, p.23).

2.7.1 Previous MMN studies

MMN is sensitive to all linguistic processes (Shtyrov & Pulvermüller, 2007, p.179; Zora et al., 2015; Zora et al., 2016a; Zora et al., 2016b). In their study, conducted with Finnish and Estonian native speaker groups, Näätänen et al. (1997) showed evidence for native-language memory traces. Their interpretation was the result of enhanced MMN responses to the language specific phoneme representations (Näätänen et al., 1997, p.433). In previous studies, comparing pseudowords to meaningful words, a word-related MMN enhancement was found (Shtyrov & Pulvermüller, 2002, p.525). This enhancement is argued to be the activation of memory traces of strongly populated neurons for the meaningful words (Shtyrov & Hauk, 2004, p.1085).

MMN’s sensitivity is not specific to just one’s native-language. Perception and thus discrimination of auditory input has also been investigated among L2 listeners. After a couple of days training, native speakers of Finnish showed evolved memory traces which was indexed by behavioral studies as well as the MMN (Tamminen et al., 2015, p.23). In an MMN study conducted with Finnish monolinguals and Swedish-Finnish bilinguals, Tamminen, Peltola, Toivonen, Kujala & Näätänen (2013) compared phonological processing between monolinguals and bilinguals. Their results showed smaller amplitudes and longer MMN latencies in bilinguals than in monolinguals, which is argued to be the result of “[e]xtensive intertwined phonological system where both languages are active all the time” (p.12).

MMN has been used in a magnetoencephalographic (MEG) study to investigate early processing for semantic and syntax conducted by Menning, Zwieterlood, Schöning, Hihn, Bölte, Dobel, Mathiak & Lütkenhöner (2005). The study was divided into two parts. In the first part a phonemic contrast was presented in a sentence ‘You have just heard lawn/giants/roses’ which is originally /ra:sen/ /ri:sen/ and /ro:sen/ in German (p.78). In the second part on the other hand the sentence was ‘The woman fertilizes the lawn/giants/roses in May’ in which on the one hand the word giant ‘riesen’ created a semantic error because one cannot fertilize giants, on the other hand the word roses ‘rosen’ created a morphosyntactic error due to German language grammar rules (p.78). The brain responses from the MEG data showed evidence for larger MMN to the semantic error than the syntactic one which interpreted by the authors as first and foremost that semantic and syntactic process were different and that “MMN reflected detection of meaningfulness” (Menning et al., 2005, p.79-80).

In the present study MMN’s sensitivity to meaning has been taken into account to investigate the brain responses by presenting the participants different types of auditory stimuli sharing a semantic relation, thus denoting same and/or similar referent.
2.8 Other ERP components of interest

2.8.1 N1 - P2 complex

In the previously mentioned ERP study of Cummings et al. (2006) (section 2.4), all of the presented stimuli elicited N1-P2 complex. The difference was that the semantically matching ES and picture stimuli pair elicited an N1-P2 complex followed by a positive wave, and the mismatching pair elicited an N1-P2 complex followed by a negative wave (p.95). The N1 component of this complex is known as an auditory evoked potential that is usually seen with a latency at around 100 ms followed by a positive P2 curve at about 175 ms (Hyde, 1997, p.282). A general view in the literature is N1-P2 amplitude increases with attended stimuli unlike MMN (Hyde, 1997, p.289).

The auditory N1 response is underlined by combination of processes ( Näätänen & Picton, 1987, p.386) that does not reflect one single event and it differs from the MMN; N1 is sensitive to an individual stimulus in contrast to MMN, which is sensitive to the relations between the present and the preceding stimuli ( Näätänen & Picton, 1987, p.389). Unlike the memory dependent ERP component MMN, N1 will be small for the stimuli with an already existing memory trace ( Näätänen, 2019, p.42) and if there is no trace there would be no MMN according to Näätänen (2019) and N1 amplitude will be a large one (p.42).

2.8.2 N200 complex

This ERP complex consists of two components one of which is N2a (MMN) and N2b ( Näätänen, Simpson & Loveless, 1982, p.87). N2b is the endogenous component which elicits when the participants attend to a stimulus unlike MMN ( Näätänen et al., 1982, p.53) The only difference is not the attended condition between N2a (MMN) and N2b. N2b has for instance a latency peaking at around 200-300 ms which is later than the N2a (MMN) component ( Näätänen, 1984, p.291) and has a posterior distribution in contrast to MMN which is known to be elicited in the fronto-central regions ( Näätänen et al., 2007, p.225).

Furthermore, N2 has been argued to be an index of inhibition during bilingual language production, but not for the comprehension (Misra, Guo, Bobb, & Kroll, 2012, p.234). However, inhibition in bilingual language processing is recommended to be further investigated (Chen, Bobb, Hoshino, & Marian, 2017, p.52). In their study Misra et al. (2012) investigated whether late but relatively proficient Chinese–English bilinguals inhibit the L1 in order to name pictures in the L2 (p.233). The first group named pictures first in their L1 and then in their L2. The second group did the opposite and first named the pictures in L2 and then in L1 (p.228). Misra et al. (2012) expected a facilitation due to priming effects if no inhibition occurred, otherwise priming was expected to be reduced or eliminated due to an inhibition effect (p.27). Their results displayed two negative activation, one at approximately around 100 ms at fronto-central regions which was followed by a positive peak, and the second at 150 ms again followed by a positive peak around 200 ms. The ERPs also displayed more positivity for L1 that was followed by the L2 condition and more negativity for L2 followed by the L1 condition. They interpreted their results as the priming occurred only for L2 and that the N2 was enhanced when bilinguals switched languages (p.234).
3 Purpose and research questions

The goal of this study is to investigate the neuronal activity when two semantically related auditory input of ES and lexical stimuli presented in an oddball paradigm by focusing on the negative event-related potential Mismatch Negativity (MMN). The assumption is that the stored language representations and the established connections in the semantic memory for those language representations would lead to distinct brain responses when comparing bilinguals to monolinguals.

3.1 Research questions and hypotheses

Question 1: Would an MMN response be elicited in a design where the two stimuli are distinct (i.e., speech sound in contrast to environment sound)?

Hypothesis 1: An MMN response would be elicited in all groups since MMN is a change dependent component and elicited by any discriminable change in the auditory stimulation (Näätänen, 1978). In the present study, the presented auditory stimuli have discriminable acoustic features.

Question 2: How would semantically related ES and word stimuli affect brain responses of monolinguals and bilinguals?

Hypothesis 2: The greatest MMN response is expected to be among the native speakers of English and Turkish for the conditions in which ES and word stimuli match in their semantic representations.

Question 3: Can bilinguals access their L2 when an auditory word stimulus is presented in their L1?

Hypothesis 3: The bilingual participants of this study will use sub-phonemic cues to activate their L2 English.
4 Method and material

4.1 Participants

The participants were native speakers of American English (n=1 female) and late Turkish-English unbalanced bilinguals (n=3 males). Their age ranged between 29 and 48 (M=38.5; SD=7.43). All of the participants signed an informed consent form prior to the experiment and filled out Edinburgh inventory (Oldfield, 1971). They were all right-handed; the laterality index for all of the four participants were 100 % for the 15 items selected and adapted from the Edinburgh Inventory. Participants reported no neurological disorders, or language related deficiencies. The Turkish native speakers also completed an online English vocabulary test LexTale (Lemhöfer & Broersma, 2012) that is designed specifically for cognitive studies to be able match the participants’ proficiency levels within the group. In the present study bilinguals’ test results vary as 80%, 48.75% and 68.75 (SD=12.9). Participants were recruited via social media and among the available acquaintances that volunteered in the experiment without any compensation.

4.2 Stimuli

Experiment contains four different stimuli, two of which are ES ‘engine’ and ‘rain’. The other two stimuli are lexical stimuli: the Turkish word ‘kar’ [kʰɐɾ], which means “snow” and the English word ‘car’ [kʰɑːɹ]. ES were selected from a number of copyright free material online which were most perceivable as such in a couple of hundred milliseconds’ duration. The lexical stimuli were recorded by a female native American English speaker in a sound-proof recording studio in the Phonetics Lab of the Department of Linguistics, Stockholm University. REAPER digital audio workstation software was used for the recordings. The Brüel & Kjær 1/2” Free-field Microphones (Type 4189) with preamplifiers (Type 2669) were connected to a Brüel & Kjær NEXUS Conditioning Amplifier (Type 2692), which in turn is connected to a Motu 8M audio interface. The presented lexical stimuli analyzed with the speech sound analysis software Praat (Boersma & Weenik, 2019). Their spectrogram and waveform images can be seen in Figure 1.

![Waveforms and spectrogram](image)

Figure 1 The waveforms and spectrogram of the speech sounds used as standard stimuli in the present study.
All stimuli were manipulated in intensity of 70 dB, which was determined based on the average intensity value of the sound file car. Moreover, all stimuli were matched in duration (600 ms). The 600 ms duration limit was set in order to get a sound effect as perceivable as possible from the ES and manipulating the speech sounds as little as possible in length, without any lexical cost. Finally, all four stimuli were processed at the beginnings and the ends of each sound file to get rid of the possible noises for a 5 ms of fade-in and fade-out effect.

4.3 Setting up the EEG

The recordings were made in the EEG lab of the Department of Linguistics, Stockholm University. The participant sat on a chair in front of a screen and the two loud speakers. The preparations began with the measurement of the participant’s head to be able find a suitable BioSemi headcap. This special cap has holes with plastic holders on it for the electrode placements. The procedure then continued with the placement of the seven external electrodes. First, to be able to catch the horizontal eye movements two electrodes were placed on the outer sides of the eyes. For the vertical eye movements two electrodes were placed on top and below of the left eye. Two mastoid electrodes were placed behind the ears and one electrode was placed on the tip of the nose as reference.

The predetermined BioSemi headcap was placed on the head. After participant’s self-identifying the center of his/her head, the electrode holes were filled with the same electrode gel that was used on the external electrodes. A 16-channel layout was used in the experiment. The electrodes that were placed on the headcap are Fp1, Fp2, F3, F4, P3, P4, C3, C4, O1, O2, T7, T8, Fz, Cz, Pz and Oz. The electrode placement on the scalp is illustrated in Figure 3.

![Figure 3 The 16-channel electrodes are arranged according to a common 10-20 system, that is a method for defining and naming electrode placement on the scalp (used with permission from BioSemi).](image)

4.4 Procedure

The experiment consisted of two 24-minute-long blocks that were designed in E-Prime 2 (Psychology Software Tools, Pittsburgh, PA, USA). Audio stimuli were presented with two loudspeakers standing
right next to a screen where a silent documentary was shown to draw the participants’ attention away from the auditory stimuli.

The stimuli were presented in a passive oddball paradigm with one standard (p=8/10) and two deviants (p=2/10) with at least two consecutive standards after a deviant. The deviants were the same in the two blocks; the sound effect of a car engine start, that is the ES ‘engine’ and the sound effect of gentle rain fall, that is the ES ‘rain’. The standard stimulus was the English word ‘car’ [kʰɑː] in the first block, and the Turkish word ‘kar’ [kʰɑɾ] in the second block. There were 1200 stimuli in total, 960 being the standards and 240 deviants that were equally divided into two in the two blocks. The stimulus onset asynchrony was set to 1200 ms and the interstimulus interval was set to 600 ms since MMN is a memory-dependent component and a longer interstimulus interval would show no MMN (Näätänen, 2007, p.134). Experiment took 48 minutes without the preparation time and the short break between the two blocks.

A schematic illustration of the auditory oddball paradigm is shown in Figure 4.

<table>
<thead>
<tr>
<th>S1</th>
<th>S1</th>
<th>D1</th>
<th>S1</th>
<th>S1</th>
<th>S1</th>
<th>S1</th>
<th>D1</th>
<th>S1</th>
<th>S1</th>
</tr>
</thead>
<tbody>
<tr>
<td>car</td>
<td>car</td>
<td>rain</td>
<td>car</td>
<td>car</td>
<td>car</td>
<td>car</td>
<td>car</td>
<td>rain</td>
<td>car</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>S2</th>
<th>S2</th>
<th>D1</th>
<th>S2</th>
<th>S2</th>
<th>S2</th>
<th>S2</th>
<th>S2</th>
<th>D1</th>
<th>S2</th>
<th>S2</th>
<th>S2</th>
</tr>
</thead>
<tbody>
<tr>
<td>kar</td>
<td>kar</td>
<td>rain</td>
<td>kar</td>
<td>kar</td>
<td>kar</td>
<td>kar</td>
<td>kar</td>
<td>rain</td>
<td>kar</td>
<td>kar</td>
<td>kar</td>
</tr>
</tbody>
</table>

Figure 4 The letter ‘S’ is for standard auditory stimuli and the letter ‘D’ is for the deviants. In this figure only a half session of one block is illustrated. The probability ratio of the standard is 80% and the deviant is 10%. In the next half of the block the deviant is the ES ‘engine’ with a 10% probability with the same standard stimulus ‘car’. The same design also applies for the second block with the Turkish standard stimuli ‘kar’.

4.5 EEG Data Analysis

The offline data analysis was conducted with MATLAB (R2019a, The Math Works Inc., Natick, Massachusetts, USA) using EEGLAB toolbox (Delorme & Makeig, 2004) following the guidelines as resented in Luck (2014) and Zora (2016). The collected data using BioSemi ActiveTwo system and ActiView acquisition software (BioSemi, Netherlands) were down sampled to the size of 256 Hz from the size of 2048 Hz in order to optimize the memory and disk storage for the analysis.

Channel locations were set in the program for the 16 electrodes that were mentioned above in section 4.3 and signals were referenced offline to the nose channel after setting up the locations. A band-pass filter was applied at 0.5–30 Hz. An automated Independent Component Analysis (ICA; Jung et al., 2000) was also carried out to remove the eye-related components from the rest of the cerebral activity.

The data were segmented into epochs between -100 to 900 ms with a -100 ms baseline. After extraction of the epochs with a baseline correction, the artifact rejection was done automatically to remove all activation that is higher than ±100 μV for all channels. The datasets were merged and then the grand average was computed for all participants and for each stimulus. A difference wave form was obtained by subtracting the standards from the deviants (deviant-minus-standard).

The statistical analysis was not carried and was not in the scope of this thesis due to the time constraints.
5 Results

5.1 ERP data for the L1 English speaker

Visual ERP analysis indicate that in the native English-speaking participant’s data, the largest neural response was for the block in which the deviant was the ES rain and the standard was the English word ‘car’. The MMN in this block was elicited with a latency of 113 ms and a peak amplitude of -2.5 followed by a positive curve at around 180 ms.

Following this, the ES ‘engine’ with the English word ‘car’ elicited the second largest MMN with an 85 ms latency and a peak amplitude at -0.9 followed by a positivity largest at 164 ms. The third MMN response was to the ES ‘engine’ when presented with the Turkish word ‘kar’ at around 93 ms and with a peak amplitude of -0.8 followed by a positivity largest at 145 ms. Finally, the smallest MMN was elicited for the ‘rain-kar’ stimuli pair at 105 ms and with a peak amplitude of -0.2. This negativity was followed by a positive curve largest at 150 ms.

The peak amplitudes and the latencies for the native English speaker are listed in Table 4 and the difference wave forms recorded from the fronto-central electrode Fz are shown in Figure 5.

Table 4 L1 English speaker ERP data gathered from the fronto-central electrode are given below. The responses in table were listed according to the sizes of the peak amplitudes beginning from the largest on the top.

<table>
<thead>
<tr>
<th>Stimuli Pair</th>
<th>Amplitude</th>
<th>Latency (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rain-minus-car</td>
<td>-2.5</td>
<td>113</td>
</tr>
<tr>
<td>engine-minus-car</td>
<td>-0.9</td>
<td>85</td>
</tr>
<tr>
<td>engine-minus-kar</td>
<td>-0.8</td>
<td>93</td>
</tr>
<tr>
<td>rain-minus-kar</td>
<td>-0.2</td>
<td>105</td>
</tr>
</tbody>
</table>

Figure 5 The ‘deviant-minus-standard’ difference wave forms are illustrated in this figure for each block in the experiment. Waveforms are illustrated in different colors: ‘rain-minus-car’ in red, ‘engine-minus-car’ in yellow, ‘rain-minus-kar’ in green and ‘engine-minus-kar’ in blue for the native English-speaking participant. Negativity in the figure is plotted upward.
5.2 ERP data for the Turkish-English bilinguals

According to the amplitudes and the latency data (Fig.5) gathered from the bilingual group, early MMN responses elicited with an intensity between -0.4 and -1.4. The largest MMN was elicited for the ES ‘engine’, in which the block standard was the English word car. This MMN response was then followed by a positive curve at around 230 ms.

The second largest MMN occurred with a 90 ms latency followed by a 218 ms positivity for the ES ‘rain’, in which the frequent standard was again the English word car with a -1.2 peak amplitude.

As for the blocks in which the standard was the Turkish word kar, the MMN was larger for the ES ‘engine’ with a peak amplitude at -1.1, than the ES ‘rain’ in which the MMN was elicited at 97 ms with -0.4 peak amplitude.

The difference wave forms recorded from the fronto-central electrode Fz for all deviant-minus-standards are illustrated in Figure 6 and the data points are listed in Table 5 for the bilingual group.

![Figure 6](image)  
Figure 6 The difference waveforms gathered from all stimuli recorded from the Fz electrode for the Turkish-English bilingual group is displayed in this figure are coded in same colors as in the English participant’s figure (Figure 5). The color codes for the difference waves are red for the rain-minus-car wave, yellow for the engine-minus-car wave, green for the rain-minus-kar wave and blue for the engine-minus-kar wave. Negativity in the figure is plotted upwards.

<table>
<thead>
<tr>
<th>Stimulus</th>
<th>Amplitude</th>
<th>Latency</th>
</tr>
</thead>
<tbody>
<tr>
<td>engine-minus-car</td>
<td>-1.4</td>
<td>105 ms</td>
</tr>
<tr>
<td>rain-minus-car</td>
<td>-1.2</td>
<td>90 ms</td>
</tr>
<tr>
<td>engine-minus-kar</td>
<td>-1.1</td>
<td>88 ms</td>
</tr>
<tr>
<td>rain-minus-kar</td>
<td>-0.4</td>
<td>97 ms</td>
</tr>
</tbody>
</table>

Table 5 The MMN latency and peak amplitude data that were gathered from the ERP data for each stimulus. In this table the listed data are gathered from the Turkish-English bilingual group.
5.3 Topographical distribution

Scalp distributions are important to compare the investigated component with previous study results. If the distributions are similar then the components can be regarded as the same (Luck, 2014, p.69). MMN has its largest amplitude over the fronto-central scalp areas (Näätänen et al., 2007, p.2549) and the distribution of the negativity is represented with the color blue in the following topography maps. All the scalp topography difference maps were scaled between [-2.5 2-5] in both groups at the peak latencies of the relevant condition that are listed in Table 4 and Table 5 and they are also given under the relevant condition in Figure 7 and Figure 8.

Scalp topography difference maps are illustrated for the native English speaker in Figure 7, and for Turkish-English bilingual group in Figure 8.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Native English</th>
<th>Turkish-English</th>
</tr>
</thead>
<tbody>
<tr>
<td>rain-minus-car</td>
<td>113 ms</td>
<td>90 ms</td>
</tr>
<tr>
<td>engine-minus-car</td>
<td>85 ms</td>
<td>105 ms</td>
</tr>
<tr>
<td>rain-minus-kar</td>
<td>105 ms</td>
<td>97 ms</td>
</tr>
<tr>
<td>engine-minus-kar</td>
<td>93 ms</td>
<td>88 ms</td>
</tr>
</tbody>
</table>

Figure 7 The scalp topography difference maps of the native English participant show fronto-central distribution which is in line with MMN except in the ‘engine-minus-car’ condition. The latencies are provided below for the relevant condition. These are the specific times in which the distribution data is gathered using EEGLAB toolbox (Delorme & Makeig, 2004). In the ‘engine-minus-car’ condition the distribution (the darker blue region in the condition) shows a parietal activation, which is not a typical scalp distribution of MMN.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Native English</th>
<th>Turkish-English</th>
</tr>
</thead>
<tbody>
<tr>
<td>rain-minus-car</td>
<td>90 ms</td>
<td>105 ms</td>
</tr>
<tr>
<td>engine-minus-car</td>
<td>105 ms</td>
<td>97 ms</td>
</tr>
<tr>
<td>rain-minus-kar</td>
<td>97 ms</td>
<td>88 ms</td>
</tr>
</tbody>
</table>

Figure 8 The scalp topography difference maps of the Turkish-English bilingual group display fronto-central activation in the ‘engine-minus-car’ and ‘engine-minus-kar’ conditions. The ‘rain-minus-car’ and ‘rain-minus-kar’ condition activations are distributed centrally, which is in line with previous MMN studies’ scalp topography maps.
6 Discussion

6.1 Result Discussion

This thesis aimed to investigate the interplay between the semantically related stimuli when presented as ES and words to monolinguals and bilinguals in a passive auditory oddball paradigm. The focus was on the event-related potential component MMN, which is known to be sensitive to long-term memory traces associated with words and sounds (Näätänen, 1997, p.432; Shtyrov & Pulvermüller, 2007, p.181) such as language-specific phonemes in one’s language would lead to larger MMN responses in contrast to non-phonemes as well as words would lead to larger MMN responses in contrast to pseudowords. Due to this sensitivity to long-term memory traces, it was assumed that the stored language representations and the established connections in the semantic memory for those language representations would lead to distinct brain responses when comparing bilinguals to monolinguals.

That is, it was assumed that for example the Turkish word ‘kar’ would lead to different brain responses between the English monolingual participant and the Turkish-English bilingual participants since the monolingual participant is expected not to have an established representation for that specific language and that Turkish word.

First and foremost, the present study results were difficult to interpret. That is, the gathered data was not enough due to lack of participants and there was not enough evidence supporting a possible explanation of the results as a whole. Therefore, they were discussed in regard to different possible reasonings individually, which in some cases became contradictory with each other. Thus, further investigation is needed for the contradictory explanations in future studies.

In the present study, the first question was asked in regard to MMN’s sensitivity to the acoustic changes between the stimuli. That is, an MMN response was expected to be elicited in all groups since MMN is a change dependent component, which is sensitive to infrequent stimuli and is elicited by any discriminable auditory change (Näätänen et al., 1989, p.217, 220). The ERP data gathered from the participants indicate that an MMN response was elicited in both Turkish-English bilinguals and the native English speaker, which is in line with the first hypothesis. This hypothesis has also been supported by the overall frontocentral scalp distribution in both groups, which is the typical distribution of MMN. However, there was one single exception in the native English speaker’s data, in the ‘engine-minus-car’ condition, which shows a parietal distribution (Figure 7). Luck (2014) mentions that the scalp distributions are important to compare the investigated component with previous study results and that the components can be identified depending on their scalp distributions in comparison to the investigated component (p.69). Previous studies of Van Petten & Rheinfelder (1995) and Cummings et al. (2006) mentioned above (section 2.4) have also used words and environmental sounds that were semantically related to them as their stimuli. Van Petten & Rheinfelder (1995) reported that the relationship of the sound to the preceding word affected the negativity wave (N400) in amplitude for the related stimuli pair in comparison to the unrelated pair (p.493). Furthermore, they reported an N1-P2 complex peaking at about 100 ms, and P2 peaking at about 190 ms both for words and for sounds (p.491). Cummings et al. (2006) also reported an N1-P2 complex for the stimuli types that they presented to their participants, however, Cummings et al. (2006) only provided the focus components latencies and not the latencies for the N1-P2 complex. In this study, for the bilingual group, negativity occurs between 88 ms and 105 ms followed by a positivity at around 220 ms. As for the native English speaker, the negativity was between 85 ms and 113ms, which was followed by a positivity between 145 ms and 180 ms. So the present study shows some prima facie similarities with Van Petten & Rheinfelder (1995) and Cummings et al. (2006) in regard to the stimuli and latencies and it is possible that one can argue for N1-P2 complex for the present study results in regard to the latencies, however, it is also important to mention that these two previous studies focus on a different ERP component, which is the N400. The so-called N400 is known to be sensitive to meaningful stimuli and gets larger in amplitude with semantic incongruencies. So, even though Van Petten & Rheinfelder (1995) and Cummings et al. (2006) had similar stimuli types their focus ERP component was not MMN. Furthermore, the distribution, which
was reported over the fronto-central or centro-parietal areas in the two studies were obtained from a latency, which is typical for the N400 component time-interval and is different and later than the investigated latency to acquire the topographic data in the present study. The scalp topography maps were acquired at the latencies provided in Figure 7 and Figure 8 (section 5.3). That is, the topographical data for the centro-parietal distribution in the previous studies showed a distribution for the N400 component and not for N1-P2. Moreover, the two studies used different experimental designs than the present study’s passive oddball paradigm. Finally, in regard to this hypothesis, it is important to keep in mind that this parietal distribution was observed only in one participant and only in one condition as mentioned earlier, which is not enough evidence for an argumentation against the MMN. Therefore, the parietal distribution in the ‘engine-minus-car’ condition in the native English speaker’s data can be argued due to an effect of an artifact, which might have been left in the data as a result of an automated artifact removal run.

As a second hypothesis, it was assumed that the largest MMN was expected among the native speakers of English and Turkish for the conditions, in which ES and word stimuli match their semantic representations. That is the MMN was expected to be the largest in the ‘car – engine’ condition for the native English speaker participant, and it was expected to be largest in the ‘kar - rain’ condition for bilinguals who have Turkish as a native language. These responses were expected so because MMN is sensitive to long-term memory traces and one gets enhanced responses for the stored language representations because of the intrinsic properties of the component (Shtyrov & Pulvermüller, 2007, p.181). That is, MMN is an index of experience-dependent memory traces (Shtyrov, Hauk & Pulvermüller, 2004, p.1083). The assumption here was that, for the native listeners, the semantic representation of the ES ‘engine’ as a deviant would lead to a stronger neuronal activity for the familiar word ‘car’ due to the established associations between the two stimuli. In addition, a further hypothesis was that an MMN response for the ‘car – engine’ was expected among bilinguals with a smaller intensity in comparison to the native English speaker for the same reasons. The gathered data from the native speaker of English, however, displayed a larger response to ‘car-rain’ condition with a -2.5 peak amplitude and 113 ms latency. Turkish-English late bilinguals, elicited the largest MMN in the ‘car-engine’ condition with a -1.4 peak amplitude and a 105 ms latency, which automatically canceled the following additional hypothesis that Turkish-English bilinguals would elicit a smaller MMN response in the ‘car-engine’ condition compared to the native English participant. Therefore, it is not possible to accept the hypothesis that, the largest MMN was expected among the native speakers of English and Turkish for the conditions, in which ES and word stimuli match their semantic representations.

Even though the hypothesis was canceled, in trying to find a reasonable explanation for the unexpected outcome is discussed in this part. MMN’s sensitivity to rare stimuli is argued to be in relation to an “expectancy violation” (Pulvermüller, Shtyrov, Hasting, & Carlyon, 2008, p.249). For example, as soon as the word ‘car’ is identified, the semantically related target word(s) are expected to be provoked (e.g., the word ‘table’ would activate the word ‘chair’ earlier than the word ‘baby’) (Hoey, 2012, p.8) as discussed earlier (Section 2.5.1). With this logic, one can argue that the word ‘car’ primes ES ‘engine’ and that the native English-speaking participant would elicit a smaller MMN for this stimuli pair. In this study, the L1 English participant has in fact elicited smaller MMN for ‘car-engine’ stimuli pair (-0.9 / 85 ms). So, one can conclude that there was no expectancy violation in this condition for the native English speaker. Therefore, it can be argued that a smaller MMN was elicited in contrast to the semantically unrelated stimuli pair. The native English speaker’s gathered data for the largest MMN also confirms this idea that there was an expectancy violation in the ‘car – rain’ condition, in which the peak amplitude and latency were -2.5 and 113 ms. The same logic also applies for the Turkish stimuli pair among the Turkish-English bilingual participants. The data gathered from the native Turkish speakers show that the smallest MMN response was to the ‘kar – rain’ stimuli pair with a -0.4 amplitude and 97 ms latency. That is, there were no expectancy violations between the two stimuli.

As mentioned earlier, the largest MMN in the bilingual group was expected in the ‘kar-rain’ condition due to the semantic relatedness and established connections. The MMN, however, was elicited largest in the ‘car-engine’ condition. This unexpected result can be argued due to an effect of the bilinguals
L2 dominance. Previous studies show evidence for positive effects on the second language with long exposure and experience. The bilingual participants of the present study reported that they actually use mostly their second language English for the longer proportion of the day since none of them actually speaks fluently the language of the country in which they reside, which is also distinct from their native language. The responses in which the lexical stimuli are English in the bilingual group are similar in amplitude (-1.2 for the ‘car-rain’ and -1.4 for the ‘car-engine’) and the discussion for the expectancy violation in the previous paragraph contradicts the results when taking L2 dominance in regard to the second stimuli pair with the English lexical stimulus, that is the ‘car-rain’ condition. So, one cannot explain the L2 dominance following the same logic for all of the stimuli. Since if L2 dominance played a significant role in the bilingual group data, one would also expect to see an expectancy violation in the ‘car-rain’ condition similar to the native English participant’s data, however, this is not the case and L2 dominance is undermined by this condition and L2 dominance is thus provides a partial explanation for a specific curve, but not all of the curves (Figure 6, section 5.2).

The final hypothesis was in relation to the specific Turkish auditory stimuli presented to the participants. In this study, the recordings were made by a native English speaker who does not speak Turkish. Furthermore, there are subtle differences (e.g., the backness or length of the vowel /a/ in [kʰɑɾ] and [kʰɑɹ]) between these interlingual homophones. Language non-selective access view claims that the two languages of a bilingual will be activated even when the stimulus is presented in one language (Schulpen et al., 2003, p.1155). And according to the bilingual memory model, BIMOLA, the heard word stimuli activate common phonological features of the bilinguals’ two languages, which then, in turn, leads to phoneme activation and then the selected language’s word activates (Schulpen et al., 2003, 1172). In this regard, it was assumed that the bilinguals of this study would use the sub-phonemic cues to activate their L2 English. So, even when they were presented with the Turkish word ‘kar’, it was expected that the English word ‘car’ would also be activated. The heard stimuli can be represented as ‘CAR-engine’ and ‘CAR-rain’ (capital letters were used to differentiate the actual English lexical stimuli from the activated English heard stimuli after the sub-phonemic cues).

Bilingual’s data, in which the presented stimuli were the Turkish word ‘kar’, the MMN responses are with a -1.1 peak amplitude and latency at 88 ms for the ‘kar-engine’ condition and -0.4 peak amplitude and latency at 97 ms for the ‘kar-rain’ condition. Firstly, taking the stored language representation assumption into account, the MMN responses in the bilingual group in which the lexical stimuli are the Turkish word ‘kar’ show that the responses were in line with the expectations and the stronger neuronal activity for the stored representations was elicited larger MMN, that is to say, CAR-engine elicited a larger MMN response in comparison to the CAR-rain. When taking the expectancy violation explanation on the other hand the same responses contradicts the stored language representation explanation and it creates a circular conflict between the two explanations. Furthermore, when taking all the stimuli in the bilingual group into account, even though the two stimuli pair, which have the same ES as deviants, elicited similar MMNs, the stimuli pair that have the ES ‘rain’ as deviants (‘CAR-rain and ‘car-rain’) are not similar in amplitudes (-1.2 and -0.4 respectively). To be able to give a robust answer, first, there should be a control group, which involves monolingual Turkish participants, and second, there should be more participants to obtain more data which is crucial in a future study. As a result, in regard to the sub-phonemic cues, the results are inconclusive to accept the third hypothesis.

6.2 Method Discussion

In the short period of time given to conduct the experiment, finding more participants was difficult and therefore, the gathered data size was not large enough to draw concrete conclusions. The total amount of participants in the present study were not matched in size or gender either; there was 1 female as a native English speaker and there were 3 male Turkish-English bilinguals. It would have been ideal if the groups were matched. Furthermore, due to the lack of participants, the English vocabulary test LexTale was not taken into account as a preselection process. So, the recommendation will be to match the variation between the bilinguals’ proficiencies for a future study to remove any other variables within the group.
The initial thesis proposal consisted of three participant groups that included also Turkish monolinguals as well as the native English speakers and Turkish-English bilinguals. However, no monolinguals were found. So, one of the hypotheses had to be removed from this thesis due to lack of Turkish monolingual participants. If not removed, the hypothesis regarding the Turkish monolinguals was that ‘Turkish monolinguals are expected to show the smallest MMN response for the block in which the standard input is the word ‘car’, since for those speakers, the English word ‘car’ and the ES ‘engine’ has not any established connections, and therefore, would not lead to an MMN as strong as the bilingual or the native English speaker group’. For the same reasons, they would show a similar MMN response for the block in which the standard stimulus is the Turkish word ‘kar’ and the deviant stimulus is the ES ‘rain’, since both are semantically related and refer to weather conditions in Turkish. This group’s data would have been crucial to be able to compare the brain responses and infer results.

One concern about the design of the present study was that both Turkish and English auditory stimuli were recorded by a native American English speaker. The stimuli could have been recorded by a nativelike bilingual speaker of Turkish-English so that there would be no phonological variation in either of the recorded stimuli. However, in this way, it became possible to address the third hypothesis. In this regard, for the analysis to be grounded on a more solid foundation, another solution could have been a lexical decision task that is conducted with native speakers prior to the EEG recordings in order to use the best lexical matching auditory stimuli for the speech sounds and a naming task for the ES.

During the experiment, participants were told to ignore the incoming auditory input. To draw their attention away from the stream of auditory stimuli they silently watched a documentary and were not given any task at all. However, an attention control could have been possible if the participants actually were given an unrelated task such as naming the total amount of animals and their kinds that they had seen on the documentary movie right after the end of the experimental session. With this way, the attention depended component N2b overlapping would be avoided (Näätänen et al., 2007, p.2548; Zora, 2016, p.55).

6.3 Further Research

Since in this thesis there were no monolinguals of Turkish, a future study involving monolinguals as a control group with more participants would make it possible to answer the addressed questions in this thesis concretely with a wider perspective. It would be much more practical to conduct a similar study with native speakers of Swedish and bilinguals who has Swedish as a second language since the experiment would take place in Sweden.

An interlingual homophone investigation can be done prior to the experiment so that the researcher would have a wider selection of stimuli. It would be ideal to add a decision task prior to the experiment to be able to choose the right combination of word and ES stimuli pair and to reduce the ambiguity and thus variability in participants’ perceptions. One would be able to find hopefully more interlingual homophone examples given the time. In this regard, a database study of interlingual homophones in Swedish and Turkish would also be created with an academic approach. An interlingual homophone pair example in Swedish and Turkish ‘nej’ and ‘ney’ meaning ‘no’ and ‘flute’ (of a specific kind) respectively sound the same in both languages.
7 Conclusion

The purpose of this study was to investigate the neuronal activity when the two auditory stimuli were presented as ES and words that were semantically related. Participants of the study were Turkish-English bilinguals and a native speaker of English. In order to investigate bilinguals’ language access, the presented lexical stimuli were chosen among interlingual homophones from Turkish and English. The homophones were the English word ‘car’ and the Turkish word ‘kar’ meaning ‘snow’. In the experimental design a passive auditory oddball paradigm was used. Lexical stimuli were standards and ES were the deviants in the design. The experiment was conducted in the EEG-lab which is situated in the Department of Linguistics (Stockholm University). The non-invasive neurophysiological technique of electroencephalography was used to investigate the brain responses across the scalp.

In the present paper, one of the focus of the investigations was the negative event-related potential MMN, since MMN is known to be sensitive to all linguistic stimuli (Shtyrov & Hauk, 2004, p.179). An MMN was elicited in both groups, confirming the first hypothesis. Furthermore, how the semantically related ES and word stimuli would affect brain responses of monolinguals and bilinguals was asked as a research question. This question was not answered with a concrete answer and the second hypothesis that, the greatest MMN response is expected to be among the native speakers of English and Turkish for the conditions in which ES and word stimuli match in their semantic representations was not confirmed. The third hypothesis that, Turkish-English bilingual participants of this study will use the sub-phonemic cues to activate their L2 English was not confirmed. The results were inconclusive in regard to bilinguals selecting the relevant language by using sub-phonemic cues, in this case English. As discussed above, due to the shortcomings in the present study, all the conclusions are only relevant for this study alone.
References


