The Influence of Language Categorization on Face Perception

Elisa Ruiz Tada

TESI DOCTORAL UPF 2018

Thesis supervisors:

Dr. Albert Costa Martinez

Dr. Cristina Baus

Department of Information and Communications Technologies (DTIC)

Biomedicine Doctoral Programme

Universitat Pompeu Fabra
Barcelona

To my family.

Acknowledgements

I would like to express my sincerest gratitude to my two supervisors. To Albert for having confidence in me and allowing me to pursue my ideas. To Cristina for sharing with me her knowledge, guidance, rigorous support and bottomless patience. I could not have done this without you.

I would also like to thank Florencia, Silvia, Xavi and Cristina, the Speech Production and Bilingualism group and my friends at Pompeu Fabra (Joanna, Kalinka, Marc Lluis, Marc G., Marco, Alice, Carlos, Gabriele, Francesca, Sara, Stefanie, Chema, Camille, Daria, Nuria, Elena, Sander, Luis, Miguel, Marti, Phillip, Elena, Nara, Rikk, Matthieu, Tristan, Katha, Alba, Judith, Alexandra, Aurelio, among others).

To my friends around the world who I cannot possibly name in one short list. Thank you for your friendship throughout the times we've shared together and the times we've been apart. I would especially like to thank the friends I met in Barcelona who were there for me during my time here (Lauren, Hanky, Meri, Nils, Hiltrud, Katya, Cristina Jara, Lena, Alfredo, Lee, Eoin).

I would like to extend a special thanks to my dearest friends in the Feminist Journal Club (Irene, Pallabi, Loreto, Edina, Adrian, Cristina, Mohit). Thank you. You were my support group, my safe space, my teachers, my favorite activists and best of all, my friends.

And finally, to my family. To Simon, Michaela (and baby!), thank you for always being there with humor and kindness. To Marina. We did it! We went through the good times and the bad with laughter and lots of tears, and I'm glad it was with you. And finally, to my parents who have supported me in more ways than I can count. 心の底から感謝しています。支えてくれてありがとう。

Abstract

As the world experiences increased international mobility, we encounter those from different racial, ethnic, and linguistic backgrounds. Therefore it is increasingly crucial to examine the ways we categorize and perceive other people. The main objective of this dissertation is to examine whether *language* is a dimension of social categorization, and whether this affects *face* perception. We also examined whether language categorization interacts with race categories, and whether this interaction affects the perception of other race faces. These issues were investigated in three studies. Firstly, we used behavioral and event-related potential techniques in an oddball paradigm to test whether language categorization affects visual face perception. We demonstrated that indeed, language is used as a social category, and this categorization affects the early stages of visual perception of faces. Secondly, we examined how language interacts with race in creating social categories. By using a popular psychological paradigm called the Memory Confusion Paradigm, we establish the robustness of language categorization, and the malleability of race categorization when crossed with different language contexts. In our final study, our aim was to understand whether native and foreign accents affect the perception and recognition of other-race faces. In summary, this dissertation has examined the effect of language on face perception, and has established that language categorization is a strong and robust effect that influences face perception.

Resumen

Como resultado del incremento en la movilidad internacional que el mundo está experimentando, es común encontrar gente de otra raza y con orígenes étnicos y lingüísticos diferentes. Así pues, es cada vez más crucial examinar cómo categorizamos y percibimos a otros. El objetivo principal de la presente tesis es examinar si el lenguaje es una dimensión de social categorización y como afecta la percepción de la cara. Además, examinamos si la categorización lingüística interactúa con las categorías raciales y si esta interacción afecta la percepción de aquellas caras con raza diferente a la nuestra. Todas estas cuestiones se investigaron en tres estudios. Primero, mediante medidas conductuales y electrofisiológicas en un paradigma de detección del cambio (*oddball paradigm*) se investigó si el lenguaje se usa como categoría social y si tal categorización afecta estadios tempranos en la percepción visual de la cara. Segundo, se examinó como el lenguaje interacciona con la raza a la hora de crear categorías sociales. Por medio del paradigma psicológico de la confusión de memoria (*Confusion Memory Paradigm*), establecimos la robustez de la categorización lingüística y la maleabilidad de la categorización racial en diferentes contextos lingüísticos. Finalmente, el último estudio tenía por objetivo entender si acentos nativos y extranjeros pueden modular la percepción y el reconocimiento de caras de otra raza. En resumen, esta tesis ha examinado el efecto de lenguaje en la percepción de la cara y ha mostrado que la categorización lingüística es un efecto fuerte y robusto que influye la percepción de la cara.

Preface

They were ready to do me violent harm, until they felt we were part of the same tribe, and then we were cool. That, and so many other smaller incidents in my life, made me realize that language, even more than color, defines who you are to people.
I became a chameleon.
My color didn't change, but I could change your perception of my color. If you spoke to me in Zulu, I replied to you in Zulu. If you spoke to me in Tswana, I replied to you in Tswana. Maybe I didn't look like you, but if I spoke like you, I was you.

-Trevor Noah
(Born a Crime: Stories from a South African Childhood)

This autobiographical quote by biracial South African comedian Trevor Noah is a personal anecdote that provides insight into the power of language and its interconnection with race.

Like Trevor Noah, as a result of a Spanish father and a Japanese mother, I have what face perception researchers call a *racially ambiguous face*. Throughout my life, I have been keenly aware of the ways my face has been perceived and more specifically, the flexibility in people's perception of my race. To Asian people, I looked Caucasian, but to Caucasian people, I looked Asian. To some, my race shifted day-to-day. I was certain my face had not altered on a daily basis, and therefore I understood that my race was in the eye of the beholder. As a multi-lingual, I have also noticed the myriad of ways that the *language* I spoke influenced the manners in which people categorize my race. When I spoke Japanese, I was more likely to be seen as Japanese, while when I spoke English, I was more likely to be seen as Caucasian. These are anecdotal stories, but as a result, I have always been motivated to understand the underpinnings between the interaction of language and face in the way we *see* someone.

Therefore, in this dissertation, I have explored the ways language is used as a social category, and how language categorization affects the way we perceive faces. To do this, we conducted three studies. First, we examined whether language is used as a social category, and whether this influences the perception of faces. We then explored how language categorization and race categorization interact when presented as simultaneously available cues. Finally, we investigated whether linguistic cues affect the perception and recognition of other-race faces.

Table of Contents

The Influence of Language Categorization on Face Perception

Chapter 1

General Introduction

Imagine walking into a bar and you notice everybody is either wearing a white shirt or a red and blue shirt. You wonder why until you hear those in the red and blue shirts cheer, and it becomes clear to you: they are watching a football game! Those in the red and blue shirts are fans of FC Barcelona, and their team has just scored. You walk to the bar and the bartender asks you if you support FC Barcelona or Real Madrid. What do you say? It might depend where you are- if you are in Madrid, you may answer Madrid, while if you are in Barcelona, you may want to answer FC Barcelona. As humans, we categorize people and ourselves into categories, and this categorization may have downstream consequences. If you answer the bartender with the correct football team, they may give you a free drink and a loud cheer: you have luckily been accepted as an in-group member, and narrowly avoided the disapproving looks of the other patrons. Now imagine the same bar during the World Cup. This time, they may all be wearing the same color shirt, cheering for the same team, shouting and sighing at the same time. As this example shows, social categories are dynamic, adaptable, and people can belong to multiple categories at once.

Some social category cues can be easily exchanged such as t-shirts denoting teams, while other social categories are much less interchangeable, such as your gender, race, or age, and other social categories may even be invisible or hidden, such as the language that someone speaks. While decades of research has been dedicated to the study of particular categories such as race (Allport, 1954; Meissner & Brigham, 2001), language has been a relatively ignored cue in social categorization (Pietraszewski & Schwartz, 2014b). Although invisible, language is a fundamental human trait, and therefore merits exploration with regards to its role in categorization.

The main goal of this dissertation is to examine the interaction of language (arguably a 'hidden' category until they start speaking) and race in creating social categories, and how that affects face perception. Therefore, in this chapter of the thesis, I will discuss and review the literature on social categorization with regards to the face and language.

1.1 Social Categorization

Everybody belongs to social categories: we can be male or female, young or old, a Real Madrid or FC Barcelona fan (or a not a sports fan at all). Categorization is a fundamental cognitive process which function is to organize, structure and process the stimuli in our environment in rapid and efficient ways (for a review see: (Bodenhausen, Kang, & Peery, 2012; Kawakami, Amodio, & Hugenberg, 2017). Social categorization is no exception in this capacity to classify, infer and easily process information. By categorizing others by kinship relations, sex, age or other social cues, it enables inferences about a range of relevant and important issues. We can infer, for example their goals and intentions (e.g. is this person a threat?), their knowledge (e.g. we may infer knowledge from older individuals), mating potential, etc. Doing so allows us to behave and interact with people from those groups in appropriate ways. However, social categorization differs from other types of categorization in one crucial way: the consideration of our *own* social category. In categorizing others, we are likely to consider our own status with respect to their category (i.e. as an in- or out-group member).

There are many ways in which previous literature have attempted to measure social categorization. One of the most common ways in which this has been done, is through recognition memory. Many studies have shown that in-group members are more easily recognized than out-group members. This has been repeatedly tested for race (the other race effect (ORE), Allport, 1954; for a review see Meissner & Brigham, 2001), gender (Palmer, Brewer, & Horry, 2013; Wright & Sladden, 2003) and age (Hills, 2012; Rhodes & Anastasi, 2012; Wiese, 2012), but also extends to seemingly arbitrary features such as university affiliation (Bernstein, Young, & Hugenberg, 2007).

Two main theories have been proposed for this phenomenon: the *perceptual-expertise* model, and the *social-categorization* model. The *perceptual-expertise* model, which is most often associated with the other-race effect, poses that people have differing experience with own- vs. other-group faces and therefore are better at recognizing own-race faces than other-race faces. There is support for this theory, as life-long learning with other-race faces can reverse the direction of the ORE (Sangrigoli, Pallier, Argenti, Ventureyra, & De Schonen, 2005). The *social-cognitive* model states that people have a tendency to think categorically about out-group members, but to individuate in-group members, and this difference leads to asymmetrical search for features in own- and other-group faces during encoding and processing, which debilitates subsequent recognition accuracy (Hugenberg, Young, Bernstein, & Sacco, 2010). These theories are not mutually exclusive, and therefore more recently, a hybrid model combining elements of both the perceptual expertise and socio-cognitive models are considered (Young, Hugenberg, Bernstein, & Sacco, 2012).

The social-cognitive aspect of the model allows for newer, previously un-unlearned categories such as university affiliation, socio-economic status, political views, etc. (i.e. while one has experience with male/female faces, and gender information is generally easily detected visibly, social categories such as university affiliation can not generally be assumed at first glance). The fact that these kinds of arbitrary and possibly shifting information affects recognition of these faces suggests that there is a top-down influence of knowledge of those categories when encoding, processing or retrieving those faces. Therefore, the next section of this dissertation will focus on the top-down influence of categories on perception, and how social categories may affect face perception.

1.2 Top-down influence of categories on perception

Social categorization occurs immediately, spontaneously and effortlessly. As soon as we see a face, perceptual characteristics of faces that are indicative of categories are extracted, leading to activation of the category prototypes. Previously, a *feed-forward* approach (see Fig. 1.1) was taken in social perception research, in which visual characteristics were thought to activate categories, which in turn activated stereotypes and attitudes (for a review see: Macrae & Bodenhausen, 2000). However, more recently, other authors have suggested a *dynamic-interactive* model (see Fig. 1.1), in which facial features feed social categories, which also activate stereotypes and attitudes, but those in turn influence those same social categories and perception of facial features in a top-down manner (for a review see: Freeman & Johnson, 2016).

Often, computer-generated morph faces are used to study top-down effects of certain features on how a face may be categorized (Freeman, Penner, Saperstein, Scheutz, & Ambady, 2011; Kim & Davis, 2010). In one study, participants categorized the race of a face from a White-to-Black face-morph continuum (i.e. computer generated faces which were designed to be within a spectrum of 100% White to 100% Black, with 13-points in between, where the middle face had 50% White and 50% Black features; (Freeman et al., 2011). These faces were presented with either a high-status attire (suit and tie) or low-status attire (blue-collar shirt). Low-status attire increased the likelihood of a face to be categorized as Black, and high-status attire increased the likelihood of the face to be categorized as White, and this effect increased as the race of the face became more ambiguous. Therefore, authors suggest that stereotypes interact with physical cues, affecting categorization of those faces in a top down manner.

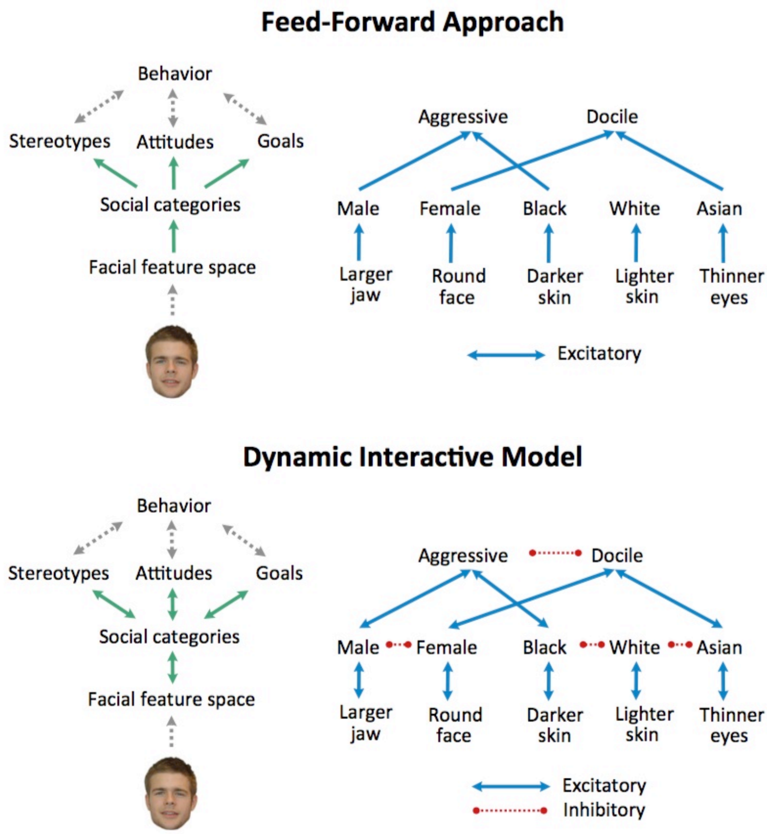


Figure 1.1 -Feed forward approach and dynamic-interactive model. Reproduced from (Freeman & Johnson, 2016).

However, there has also been an increased interest in whether cognitive mechanisms truly affect *visual perception* in a top-down manner (Collins & Olson, 2014; Firestone & Scholl, 2015; Gilbert

& Li, 2013; Vetter & Newen, 2014). Firestone and Scholl (2015) list several pitfalls to such behavioral studies, such as the difference between changes in perception and changes in judgment. They state that in many studies, especially those which rely on participants' reporting, it is difficult to disentangle whether participants have experienced a shift in *visual perception* or a shift in *judgment*. They use a shoe as an example: while one can perceive the color and size of a shoe, one can only judge (or infer) whether the shoe is expensive or not. If researchers were to ask participants the price of the shoe, they will be measuring a judgment by the participant, whereas if researchers were to ask them of the color of the shoe, participants will be reporting both their perception and judgment of the color. In a similar way, it is difficult to disentangle whether stereotypes (such as attire) have truly affected a participant's visual perception of those faces to be Black or White, or whether they have affected their judgment.

Therefore, apart from behavioral studies, neurocognitive measurements such as event-related potentials (ERP) have been widespread in examining top-down effects of categories on low-level visual perception (Holmes, Franklin, Clifford, & Davies, 2009; Thierry, Athanasopoulos, Wiggett, Dering, & Kuipers, 2009; Yu, Li, Mo, & Mo, 2017). The largest body of evidence on the top-down effect of categories on visual perception comes from color perception studies (Athanasopoulos, Dering, Wiggett, Kuipers, & Thierry, 2010; Clifford et al., 2012; Thierry et al., 2009), but this methodology has also been extended to examine the effect of categories on the visual perception of faces as well (Czigler, 2014; Kecskés-Kovács, Sulykos, & Czigler, 2013; Stefanics, Csukly, Komlósi, Czobor, & Czigler, 2012).

Therefore, in this dissertation, we have used event-related potentials (ERPs) to measure whether categorization of faces have affected the *visual perception* of those faces (Chapter 2), and have used implicit methods to measure social categorization (Chapter 3), as well as behavioral methods to test whether visual perception is affected by categorization (Chapter 4).

1.3 Language as a social category

Throughout history, there are examples of language used as a marker of whether someone was an in- or out-group member, so much so there is a term for this: a *shibboleth*. For example, during World War II, American soldiers used the word *lollapalooza* as a shibboleth to identify Japanese-Americans from Japanese spies. This was based on the premise that Japanese speakers could not distinguish or pronounce the L/R distinction. If the unidentified person were to mispronounce the term, it revealed them to be either native English speakers and therefore Japanese-American, or non-native speakers of English, and thus spies (Gramling, 1942). In addition, strictly enforced single-language policies are also used by many authoritarian governments (Liu, 2017) with the intention to eliminate minority language identities, and to dominate a people via a single language identity. During Franco's dictatorship, for example, Castilian Spanish was enforced throughout Spain. The many historic examples of such shibboleths and language policies embody the strength of language and accents as an identifier of social categories.

The Ethnolinguistic Identity Theory (ELIT; Giles & Johnson, 1987) states that language is one of the most, if not *the* most important aspect of social categorization of self and others. Using language as a social category emerges early in childhood (for a review see: Liberman, Woodward, & Kinzler, 2017) and continues throughout life.

In a famous study, Kinzler and colleagues (Kinzler, Dupoux, & Spelke, 2007) allowed five-year old monolingual children the choice to be friends with one of two children: a native language speaker (English), or a foreign language speaker (French). Children chose the native language speaker, thus the authors concluded that language influenced explicit social preference in childhood. In the same study, six-month-old, monolingual, American infants were shown two adult women speaking in either their native language (English) or a foreign language (Spanish), and in a subsequent silent test phase, those infants showed a preference (looked longer) for the English speaker (Kinzler et al., 2007). Therefore, language clearly affects explicit social preferences in five-year-olds, and looking-time preferences as early as six-month-old infants.

Evolutionary psychologists also point to the evolutionary origins and validity of having accent as a *dedicated* dimension of social categorization (Pietraszewski & Schwartz, 2014a, 2014b), i.e. a cognitive system dedicated to differentiating language from non-language and one language from another. They argue, that even when considering the limited scale of their travel (including travel by foot), they were very likely to come across linguistic variation such as different accents or languages (Chapais, 2010; Nettle, 1999; Pietraszewski & Schwartz, 2014a, 2014b). Therefore, they believe that linguistic diversity was a recurrent feature of human ancestral environments.

In more recent adult populations, Stevenage, Clarke, and McNeill (2012) found that participants were able to recognize voices from their own-accent (English vs. Scottish) better than voices from the other-accent. This phenomenon was named the Other-Accent Effect (OAE), due to its similarity with the Other-Race Effect (ORE) (Allport, 1954) in which recognition for the group with the 'other' characteristic is decreased. Recent research has argued that the "big three" social categories, which will be discussed in the next section, are incomplete, and that language could be considered the fourth dimension in social categorization (Kinzler, Shutts, & Correll, 2010).

Accents and languages are both a variation of linguistic diversity in a globalized environment. All languages and all accents inform us of the geographical origin of the speaker. However, accents and languages can cue us to different social information. Accents such as regional dialects and foreign accents may give different information as to the mastery of that language by that speaker (i.e. regional dialects inform us they have mastered that language, whereas those with foreign-accents may have not), or socio-economic class (Fuertes, Gottdiener, Martin, Gilbert, & Giles, 2012), among others. Someone speaking to us in our native or non-native language may inform us of their origin, but we can assume they have mastered their language. In addition, this may require a different cognitive load as compared to someone speaking in our own-language, but with a different regional dialect. Within this dissertation, I have used both languages (Spanish and English) as a social category (Chapters 2 and 3) as well as different accents (native and foreign-accented Spanish; Chapter 4). While the subtleties of information retrieved from accents and languages may be relevant in other studies, in this dissertation, their purpose was to create two linguistic social categories, and therefore we assume that their effects may be similar.

1.4 Race as one of the 'big three' social categories

If you were asked to recall someone, it is said that you are most likely to remember, if nothing else, the 'big three' social categories they belonged to: their age, their gender and their race (Allport, 1954; Hills, 2012; Meissner & Brigham, 2001; Palmer et al., 2013; Rhodes & Anastasi, 2012; Wright & Sladden, 2003). Among the 'big three' categories, race has been studied for decades (Allport, 1954), and had been regarded as an automatic, spontaneous, inevitable category that was encoded once we met someone new (Cosmides, Tooby, & Kurzban, 2003).

Decades of research considered that encoding of race was inevitable because people categorized others based on race even in the absence of instructions (Taylor, Fiske, Etcoff, & Ruderman, 1978), and because attempts to increase or decrease race encoding had failed (Bornstein, Laub, Meissner, & Susa, 2013; Devos & Ma, 2008; Hehman, Stanley, Gaertner, & Simons, 2011; Hewstone, Hantzi, & Johnston, 1991; Stangor, Lynch, Duan, & Glass, 1992; Taylor et al., 1978). More recently however, studies have found situations in which race categorization is reduced or altered (Freeman et al., 2011; Karl Christoph Klauer, Hölzenbein, Calanchini, & Sherman, 2014; Kurzban, Tooby, & Cosmides, 2001; Noymer, Penner, & Saperstein, 2011; Penner & Saperstein, 2008; Pietraszewski, 2016; Pietraszewski, Cosmides, & Tooby, 2014).

Recently, evolutionary psychologists studying race have suggested that race is not an inevitably activated social category, but a byproduct of coalitional psychology. They propose that human ancestors did not meet others from another race, and would therefore not have a cognitive system dedicated to the encoding of race. However, race in modern society is a cue for many contemporary social boundaries, and therefore feed the mechanism created to trace coalitional groups (Kurzban et al., 2001; Pietraszewski, 2016; Pietraszewski et al., 2014). They have found that when race was crossed with coalitional information (charity group information), categorization by race was reduced when compared to categorization strength when race was the only cue. Therefore, these authors suggest that race is a malleable cue, which can become less important in the face of other coalitional group information.

If race is indeed a malleable social category, we are left with the question as to which categories make it so. Therefore, goal of this dissertation is to look at race as a social category, and how it may be affected by other cues, namely language. Is the perception of someone's race affected by language, the same way race categorization is affected by coalitional psychology? Or will racial categorization remain robust and unaffected by language? These are some of the questions we aimed to answer.

1.5 Multiple Group Membership

Nobody belongs to just one social category. We are not just female, or just Asian, or just a student of Pompeu Fabra University. We can be all of those categories simultaneously, and more. Categorizing others simply as in- or out-group members may be complicated when multiple group memberships are available. On which dimension do we categorize them relative to ourselves? Globalization has made it increasingly likely that we meet individuals of many other social groups, whether it be geographical origins, age, race, sex, accent, language or any other dimension, and they may also interact to create complicated in/out-group dimensions across several categories.

Immigration has made it more likely that someone from another race for example, could, in fact, speak our native language (e.g. second generation immigrants or adoptees), while someone from the same race could speak with a foreign accent (e.g. foreigners). This contemporary environment makes the enquiry of the interaction of many social categories increasingly necessary to study.

Studies using multiple categories typically use two dimensions causing targets to fall within four types of memberships: double in-group, in-group/out-group, out-group/in-group and double out-group members. Patterns of evaluation of such members across studies are inconclusive and mixed, but generally fall into one of three main patterns (for a review see: Crisp & Hewstone, 2007). Some studies have found an *additive* pattern, where double in-group members are preferred over partial in-group members which are in turn preferred over double out-group members (Crisp & Hewstone, 1999, 2007). Other studies have found a *social exclusion* pattern in which double in-group members are preferred over all other groups (Kenworthy, Canales, Weaver, & Miller, 2003; Shriver, Young, Hugenberg, Bernstein, & Lanter, 2008). Finally, there are studies that have found a *partial in-group* pattern in which a second category matters only for out-group members on the first category dimension (Hehman et al., 2011; Nier et al., 2001). For example, university affiliation did not affect recognition of White faces (by White participants), but Black faces from the same university as the participants were more likely to be remembered than Black faces from another university (Hehman et al., 2011). It is likely that different social categories do not act in the same ways in affecting evaluation. That is, university affiliation may be a much more arbitrary feature compared to sex or race, and therefore different combinations of dimensions may result in different patterns of evaluation. Simply, crossing race with university affiliation may create different patterns of social categorization when compared to race with sex. Therefore, it is necessary to further investigate the ways in which different categories interact with each other to replicate a more accurate, multi-dimensional, social reality.

Therefore, in this dissertation, we examined the interaction of *language* and face race in creating social categories (Chapter 3) and how a native or foreign accent affects how other-race (Asian) faces are perceived (Chapter 4). In those studies, we manipulate in/out-group categories by language and race, to create double in-group members, double out-group members, and those who are in-group in one dimension but out-group in another (for e.g. a Caucasian face (racial in-group) who speaks the participants' L2 (language out-group)). Therefore, in the next section, we discuss previous literature examining the interaction of language and facial features in person perception.

1.6 Race and Language

Surprisingly, there are no studies directly looking at the interaction of race and language in creating social categorization when both cues are available. That is, when we meet own-race/own-language individuals, own-race/different-language individuals, different-race/own-language individuals and different-race/different-language individuals, do we use racial categories or language categories? Modern society makes it increasingly likely to meet people from any race speaking in your own language, as well as someone from the same-race speaking in a different language. Therefore, this question is increasingly becoming more relevant to discuss.

The only study which directly crossed language and race (own-race, other-language faces vs. other-race, own-language faces) have been conducted on children (Kinzler, Shutts, DeJesus, & Spelke, 2009). When five-year olds were given the choice between a playmate from their own race with a foreign-accent, or a playmate from another race with a native-accent, the children chose the latter. The authors concluded that languages and accents are important features in guiding social preferences in the case of a child's choice for playmates. Note however, that children were not given the option to play with same-race/same-accent (double in-group) playmates or different-race/different-accent (double out-group) playmates. Perhaps if children were given the choice of same-race/same-accent playmates, results may have differed. Therefore, it is also important to test different combinations of membership together, to examine the myriad of ways that categories interact to shape the way we interact with others.

Studying the emergence of racial and linguistic categorization in children is important for the understanding of when and how social categorization develops and is structured throughout a lifetime. However, it is also important to note that social categories and social interaction is learned throughout life. Explicitly asking an adult who they would prefer to interact with in a racial context will most likely influence participants to give the politically correct answer in fear of being seen as racist or discriminatory. Therefore, research conducted on adults must be done implicitly and without the realization of the participant.

No such study directly crossing language and race have been conducted on adults. One study in adults compared accents with *'ethnic looks'* (Rakić, Steffens, & Mummendey, 2011). Rakić et al., (2011) crossed accents (native and foreign) with a person's 'ethnic look' (typical German looks vs. typical Italian looks). They first tested whether participants categorized faces to the same degree based on typical ethnic looks (typical German or typical Italian faces) and accents (native- or foreign-accented German) and found that both looks and accents trigger social categorization. In a second experiment, the authors crossed the looks and accents to create four groups of faces (typical German looking face with native-accented German, typical German looking face with an Italian accent, typical Italian face with a native-accented German, and typical Italian face with an Italian accent) to test the influence of looks and accents on ethnic categorization when combined. Results showed that participants had better memory for accents than for looks, allowing authors to conclude that accents provide more relevant information in categorization when compared to 'ethnic looks'. This study is often cited as supporting the idea that visual cues are of minor relevance in social categorization when compared to linguistic cues. However, both faces used in this study were from the same Caucasian race (German and Italian). While there may be visible, stereotype differences between Germans and Italians, such ethnic differences seem hardly enough to conclude that accent or language information surpasses *race,* a much more visually salient and less ambiguous cue than 'ethnic cues', in social categorization. Findings from Rakić et al. (2011) is more in line with studies that used computer-generated face-morphs in which some faces are racially ambiguous, and therefore may be interpreted differently according to its surrounding context (e.g. social status in Freeman et al., 2011).

Therefore, this dissertation will study the interaction of race and language in implicit social categorization and person perception in adults. In the third chapter of this thesis, we examined how social categories are created when different language cues are available in own- and other-race contexts, as well as how race cues are used in native (L1) and foreign (L2) language contexts. Then, we crossed both cues to see how language and race categorization influence each other. In the fourth chapter of this dissertation, we explored whether native- or foreign-accent cues influence the visual perception and recognition of other-race (Asian) faces.

1.7 Scope of the dissertation

The main aim of this dissertation is to examine the role of language in social categorization, and its influence on the face. We have also considered the ways language interacts with race in creating social categories, and how that also influences other-race face perception.

In Chapter 2, we explored how the language that a person speaks affects early visual perception. We examined this by assigning faces as Spanish or English speakers, and measured ERPs during an oddball paradigm experiment to investigate whether the language those faces accompanied earlier affected early visual perception of those faces. In doing so, we examined the top-down influence of language as a social category on the visual perception of faces.

In Chapter 3, we investigated the interaction of language and a face's race in creating social categories. Using the memory confusion paradigm, previous studies have established that accents are a dimension of social categorization (Pietraszewski & Schwartz, 2014b, 2014a; Rakić et al., 2011), and we build on these studies to examine how language may also act as a social category. Previous studies have also looked at race as a social category (Pietraszewski et al., 2014). However, no study has examined social categorization when both cues are present in a multi-dimensional situation. That is, when we are presented with Caucasians who speak Spanish, Caucasians who speak English, Black people who speak Spanish and Black people who speak English, do we categorize these faces by language or by race, or both? By doing so, we examine ways in which faces and languages cues interact in creating social categories of faces.

Finally, in Chapter 4, we explore how language may affect the way we perceive other race faces. While Chapter 3 examines the interaction of these two categories in creating social categories, in this chapter, we examine whether other-race face perception is affected when they either speak in a native or non-native accent, cueing in/out-group membership. In this study, we used Asian female faces and Spanish sentences with a native or foreign accent, and created three groups of faces: Caucasian females who speak Spanish with a native accent, Asian faces who also speak with a native Spanish accent, and Asian faces who speak with a foreign accent. By doing so, we examined whether the Asian face with a native accent (partial in/out-group) would be comparable to an Asian with a foreign accent (double out-group that matches in race) or a Caucasian with a native accent (double in-group that matches in accent). In addition, we conducted a control experiment in which the same faces were trained with Spanish or Chinese flags instead of voices. By doing so, we aimed to test whether linguistic cues gave rise to a different as a different type of social information.

Through these three studies, we examined the ways language influence face perception, and its role in creating social categories when crossed with race.

Chapter 2

Top-down effects of language categorization on early visual perception of faces.

2.1 Introduction

The language of a speaker is one of the most important ways in which we view and categorize ourselves and others (GILES & Johnson, 1987; Pietraszewski & Schwartz, 2014a, 2014b; Rakić et al., 2011). It affects the way we identify, stereotype, trust or even remember them (Baus, Bas, Calabria, & Costa, 2017; Hansen, Rakić, & Steffens, 2017; Kinzler, Corriveau, & Harris, 2011; Lambert, Hodgson, Gardner, & Fillenbaum, 1960; Lev-Ari & Keysar, 2010). While it is important to consider how language affects the way we treat and behave around someone, little is known about the top-down influence of language categorization on early stages of face processing. The question we aim to address here is whether the language of a speaker modulates how we perceive their face.

There have recently been an increased interest in how or *if* cognitive mechanisms affect general visual perception in a top-down manner (Collins & Olson, 2014; Firestone & Scholl, 2015; Freeman & Johnson, 2016; Gilbert & Li, 2013; Vetter & Newen, 2014). Many studies have looked at how different cognitive mechanisms such as the perceiver's motivation, physical state, semantic knowledge or categorization can alter the way we see faces, colors, cars, distances, size, etc. For example, a computer generated face with ambiguous race features (morphed to be 50% White and 50% Black) was more likely to be categorized as White when it came with a higher socio-economic outfit (suit and tie) whereas it was more likely to be categorized as Black when it came with a lower socio-economic outfit (blue-collar shirt) (Freeman et al., 2011). These studies suggest that categorical learning modifies perceptual representation, or at the very least, attention or decision processes.

However, when reviewing literature on top-down influences on visual perception, several authors have pointed out that it is necessary to consider whether cognitive mechanisms truly affect *visual perception*, or instead, solely affect *judgment* of the stimuli (Firestone & Scholl, 2014, 2015). That is, often times it is difficult to tease apart whether a behavioral measure is a true indication of an alteration of visual perception or a shift in judgment. Therefore in our study, we applied event-related potential (ERP) techniques to measure early stages of visual perception.

The oddball paradigm in ERP studies has been extensively used to examine visual perception, and has also been used in categories of faces such as gender or emotion (Kecskés-Kovács et al., 2013; Kovács-Bálint, Stefanics, Trunk, & Hernádi, 2014; Kreegipuu et al., 2013; Li, Lu, Sun, Gao, & Zhao, 2012; Stefanics et al., 2012; Zhao & Li, 2006). In an oddball paradigm, a string of standard stimuli are presented, and is interrupted by a deviant stimulus. A visual mismatch negativity (vMMN) is elicited when participants detect changes in the stimuli flow, and therefore is used as a measure of visual perception and deviant detection (Czigler, 2014). Kecskés-Kovács and colleagues (Kecskés-Kovács et al., 2013) designed an oddball-paradigm testing whether a face's gender category elicited the vMMN when disrupting the continuously presented gender category. Either several female faces were presented as the standard sequence with an infrequent male deviant face, or vice versa. They found deviancy effects in which deviant stimuli elicited a more negative amplitude than the standard. The authors argued that the gender category of a face is automatically detected and that the violation of the category is reflected in the vMMN.

Other studies have tested facial emotions in oddball paradigms to explore whether emotional categories are automatically detected when unattended (Kovács-Bálint et al., 2014; Kreegipuu et al., 2013; Li et al., 2012; Stefanics et al., 2012). In one study, happy and fearful facial emotions were tested (Stefanics et al., 2012). Standard faces were of different identities and both genders, but all expressed one emotional category (happy or fearful). Deviant faces expressed the other emotion. Authors found a vMMN effect in the 170-360ms time range, suggesting that even when faces and their gender are varied and unattended, the emotional category of the face was automatically represented in standards and detected when a regularity from that category emerged.

These studies on the effect of facial categories on visual perception demonstrate primarily that even while there is a variation of stimuli within the standard presentation (different individuals of the same gender, or different identities with the same expression), the deviant category was detected and elicited a vMMN. In addition, these studies allow for the possibility to test categorical effects on face perception using the oddball paradigm.

One caveat of these oddball paradigm studies conducted on faces is that the category in question is presented in a bottom-up fashion. That is to say, that gender or emotions are visibly detectable categories available within the stimuli. Our goal, therefore, was to examine whether *language* categories affect visual perception of faces in a top-down manner. That is, whether the categorization of people by the language they speak affects the way we perceive them. The above mentioned studies have looked at previously established categories such as gender or emotions, which are easily detected from the stimuli themselves. Our study differs in that the categories were created implicitly in the lab, and the faces on their own do not generate the categories (such as gender). By showing faces accompanied by one of two languages, we first examined whether faces were categorized by language, and later explored whether the newly formed language categories affect visual perception of those faces.

The largest body of evidence on effects of categorization on early visual processing comes from studies on color perception, examining whether the way we categorize colors affects early visual perception (Athanasopoulos et al., 2010; Mo, Xu, Kay, & Tan, 2011; Thierry et al., 2009). Thierry and colleagues (Thierry et al., 2009) examined whether language-specific color terminology affected color perception. The Greek language has two different words to distinguish light and dark blue, while English only has one word for both shades of blue. Using the oddball paradigm, they explored whether there was a difference between the detection of deviance of light and dark blues in Greek speakers who have two categories of blues (*ble* and *ghalazio*, meaning dark and light blue respectively), and compared them to English speakers who only have one word for the two blue colors (blue). They found that the vMMN was similar for blue and green deviants for English speakers, while the vMMN was much larger for the blues for Greek speakers. Only Greek speakers showed a vMMN in response to the two blues. This means the Greek participants were able to implicitly discriminate the difference between the two blue colors due to having specific terminology in their language, and that the possession of these terms affected early visual perception (see Mo et al., (2011), for the same result with Chinese observers).

In another experiment, Clifford et al. (2012) used colors to examine whether *newly learned* color categories affected early visual processing. They trained one group of participants to learn new categories for green, while another group did not. They then examined whether the newly learned color categories early or late stages of visual processing during an oddball paradigm, and whether this differed between the group that had learned the new categories, and the group that had not. Category effects were only found in the trained group in what they call the post-perceptual stages of visual processing. Due to their findings, they argue that newly learned color categories affect cognitive mechanisms, but that they are independent from early perceptual processes.

Most recently, a study examined the effect of newly learned categories on face perception (Yu et al., 2017) in a similar design. Faces were trained to be one of two groups, and later examined in an oddball paradigm. The authors found similar effects of categories in the post-perceptual time windows as Clifford and colleagues, (2012), but in addition, they also found that category differences were found in early stages of face perception.

In sum, the aim of this study was first to examine whether faces were categorized by language, and second, whether low-level perceptual processing of faces was affected by the newly created *language* category (Spanish or English). First, in order to determine whether language is used as a social category, the memory confusion paradigm (MCP, also known as the 'Who Said What' paradigm; Klauer & Wegener, 1998; Taylor, Fiske, Etcoff, & Ruderman, 1978) was used to show participants the faces with a language, and determine whether they had been categorized. [The next chapter of this thesis (Chapter 3) will explain this paradigm in further detail.] We showed participants eight faces: four of which accompanied Spanish (L1) and the other half with English (L2). As with Pietraszewski & Schwartz (2014a, 2014b), we used the MCP, which is a standard way of measuring implicit social categorization. The reasoning behind this paradigm, is that if a particular feature of a person - such as the language they speak - is a basis for categorization, then people who share those characteristics are more easily confused among each other during recognition (i.e.- Spanish faces will be confused with other Spanish faces, while English faces will be confused with other English faces). This can happen without the participants' conscious awareness of this happening, and therefore reveals fundamental, implicit and automatic categorization processes.

The second aim of the present study was to test whether language categorization affected visual perception of those faces, and to do this, we used an oddball paradigm. In the oddball paradigm, three faces from one language group were designated as standards, disrupted by two types of deviants: one face from the same language group as the standards, and the other from the other language group. We used the reverse control procedure in order to test conditions in which Spanish faces were the standards or English faces represented the standards. By doing so, we explored whether deviancy from the *category* of the standard affected early stages of visual processing.

Therefore, in sum, we first tested whether participants categorized faces according to two language groups, and later examined whether these implicitly created language categories affected early visual perception of those faces in an oddball paradigm.

2.2 Methods

In order to make sure that indeed, faces were categorized by language, the memory confusion paradigm (MCP) was used following the methods of Pietraszewski & Schwartz (2014a, 2014b). Then, an original oddball paradigm test was designed to see whether the categorization of faces by language affects early perceptual processing.

2.2.1 Participants

A total of 57 participants participated in the study from the database at Pompeu Fabra University who were Spanish dominant speakers who spoke English as a foreign language. Thirty-two of them participated in the first original design of the experiment (19 male), while another 25 (10 male) participated in the second design (details of both paradigms explained further below). Due to technical and/or artifact rejection, data from 26 in the first, and 21 from the second experiment was analyzed, for a total of 47 participants. A comparison of these two groups showed no statistical differences, and therefore all analyses are on all participants. An English listening exam was done while the EEG materials were being set up for each participant to determine the English level of the participants. Out of 7 correct answers, participants made an average of 5.4 (SD=1.6) correct answers (correct answers ranged from 1 to 7), showing that they were mid-proficiency English listeners.

2.2.2 Stimuli

Grey scale photos of eight Caucasian males with neutral expressions were selected from free, electronic databases which were downloaded from the web. They did not have any easily distinguishable facial characteristic (such as a mole) on their face. They all had dark hair and eyes, so that they would be convincing as both a Spanish or English speaker. Adobe Photoshop (version CS3 for Windows) was used to isolate their face and hair from the background, and the photograph was digitally edited in order to make their hairstyles as similar as possible.

Twenty-four neutral, non-autobiographical sentences were created in Spanish, as well as its English translation of each (e.g.: El libro tiene cien páginas. [English: The book has a hundred pages.] For the full list see the Appendix). On average, Spanish phrases were 4.9 words in length, while English phrases were 5.1 words in length. No statistical difference was found between sentence lengths of the two languages ($t(23)= -1.072$, $p=0.3$). Both Spanish and English phrases ranged from four to seven words per phrase. Phrases were recorded and edited using Audacity (v 2.0.3) from four native Spanish speakers for the Spanish sentences, and four native English speakers for the English sentences.

2.2.3 Experimental Design

Participants completed two experimental sections: the Memory Confusion Paradigm (MCP), which measured social categorization, and the oddball paradigm, in which we tested whether social categorization affected early visual perception as reflected in event related potentials (ERPs). The experimental design of both paradigms will be explained further below. Participants were seated in front of a computer screen in an ERP cabin after instructions were given and consent forms were signed. The experiment was presented to the participants using the software E-prime 2.0 (Schneider, Eschman, & Zuccolotto, 2002). Participants were instructed to minimize blinking, eye movement and to stay still as possible during the experiment.

Memory Confusion Paradigm

The memory confusion paradigm or the ‘who said what’ paradigm consists of two main sections. The first is the encoding phase and the second is the recognition phase. During the encoding phase, individual photos were presented on the screen while the phrases were presented auditorily through headphones. Participants were told to form impressions of the speakers while they watched and listened, as they would be asked questions later regarding this portion of the experiment.

Each of the eight faces appeared three times during the encoding phase, for a total of 24 presentations. The three presentations of each face had a different sentence, but voices were kept the same. Simply put, each face always had the same voice, but spoke three different sentences. Photos and audio were presented simultaneously, and the photo remained on screen for a total of 4010ms. This time was selected to allow the face to appear an extra 2000ms after the longest sentence, which was 2010ms in length. A blank grey slide appeared for 200ms between each photograph (for the procedure see figure 2.1).

During the Recognition phase, all eight photos were presented in one screen, numbered 1 to 8. At the same time, the same sentences were presented auditorily. The participant had to decide which of the eight faces accompanied that sentence in the Encoding phase by pressing the corresponding number on the keyboard (i.e.- who said that sentence). The eight faces remained on screen until the participant responded. After the response, there was a blank screen for 1000ms. This was done until all 24 sentences that the participants heard in the Encoding phase were presented. Eighteen lists were created to counterbalance the face, sentence and language. Therefore, all faces accompanied every sentence in both languages across participants.

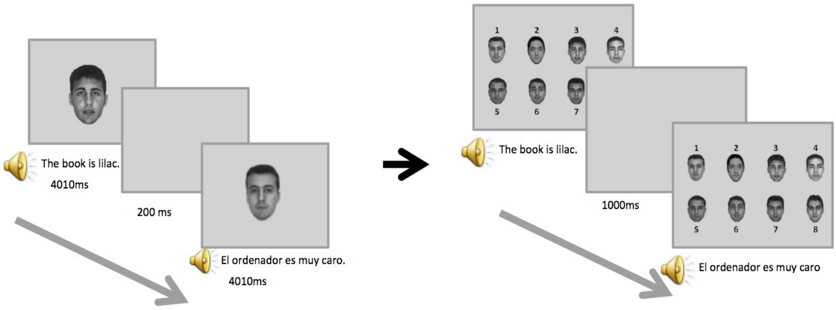


Figure 2.1 The procedure of the Memory Confusion Paradigm (MCP). This diagram shows the MCP. The very left is the encoding phase, where faces were presented with audio. Then participants played one minute of Tetris as a distractor phase, and then finally did the recognition phase.

Oddball paradigm phase

Faces that had previously been categorized as Spanish or English speakers were shown to the participants in a fast and consecutive manner, where certain faces were presented as standards (frequently seen faces) and other faces as deviants (less frequent stimuli).

Participants completed eight blocks, each with 892 face presentations (average of 748 standards, 48 deviant-within, 48 deviant-between and 48 framed faces). Four blocks had Spanish (L1) faces as the standard condition and the other four had English (L2) faces as the standards. That is, the faces that appeared the most frequently within a block (called "standards" in oddball paradigm experiments) were from one of the two language groups. Within a block, three of the four faces from one language group were repeated as standard faces and the deviant condition was either the fourth face from the same language group (deviant from within a language category) or a deviant from the other language group (deviant from between language categories). All faces were presented for 300ms, with an inter-stimulus interval of 200ms. Within a block, there were a total of 96 deviants. Forty eight of those deviants were from the same language category as the standard (deviant within language), as well as 12 presentations for each of the four faces from the other language category (deviant between language). In order to have each of the eight faces in all three conditions (standard, deviant within language and deviant between language), there were eight blocks to counterbalance the combination of conditions. Therefore, after completing all eight blocks, each face was presented as all three conditions.

In a second control experiment, in order to control for the difference in frequency of the two deviant types, deviant frequencies were controlled for. That is, in the original design described above, both deviant types had a total presentation of 48 times, but deviant within-language was one face presented as the deviant 48 times, while the deviants between-language were four faces repeated 12 times each. In this control experiment, the standards and deviant-within language remained the same, but one face was selected from the between language group and presented 48 times. By doing so, we eliminated the possibility that participants were detecting the difference in the frequency of the presentation of the two deviant types.

In order to avoid the participant from predicting when the deviant appeared, deviants appeared after seven, eight or nine standards, with 31 catch trials per block where a deviant did not appear. Deviants were split equally in these three positions so that a third of deviants appeared at each of these three positions.

As in all oddball paradigms, the participant had a task in which they would not be attending to the individual faces being presented on the screen. Therefore, the task was to press a key when the face was framed by a simple line box, which appeared every five to ten standards. These faces were presented until the participant responded, but for a maximum of 1000ms, in order to allow the participant to blink.

Recognition Phase 2

After the oddball paradigm, the same recognition test was done to see if the participants had retained the language categories throughout the oddball paradigm test. The procedure was exactly the same as the first recognition test, however, it was not preceded by a second encoding phase.

2.2.4 EEG recording and analysis

During the experiments, EEG signals were recorded using Brain Vision from 32 cap-mounted electrodes (ActiCap) organized according to the Standard International 10-20 system and referenced to the left mastoid. The sampling rate was 500 Hz, with a sampling interval of 2000 µS. The impedance was kept under 15k Ω.

Horizontal EOG was recorded from electrodes attached to both the left and right outer canthi of the eyes, and vertical movement was monitored with an electrode placed below the right eye.

ERP analysis was done using MATLAB (R2010b version 7.11.0.584) and the Eeglab Toolbox. The data went through an offline filter of 0.1 to 30Hz, and was re-referenced to the average of the two mastoids. An ICA filter was run, and eye movement components were removed. Then, ERP waves of the different trials were averaged per participant, with an epoch of -50ms to 450ms, and baseline correction of -50ms. Artifacts were rejected with a threshold of -75 to 75 µV. In cases of a malfunctioning electrode, those electrodes were interpolated from surrounding electrodes.

EEG recordings were locked to the onset of every face presentation (standards, deviants within language, deviants between language). For the analyses, only standards which were within four faces before the deviant were included, but only those that did not include a target face within those four faces. Only deviants following those selected standards were selected. This was done because target faces were instructed to be used also as a moment participants were allowed to blink, and therefore, eye movement contaminated the ERP signals. For this reason, only deviants and standards that appeared at least more than 2400ms post target were selected. An average number of 515.5 (SD=24.5) standards, 258.9 (SD=11.9) deviant-withins and 234.1 (SD=13.1) deviant-betweens were analyzed per participant.

For the statistical analyses, first, language conditions were collapsed to create three stimuli types (standard, deviant-within and deviant-within). Later, we split the conditions by the linguistic context they appeared in. That is, in Spanish context, Spanish standard, Spanish deviant-within and English deviant-between faces were analyzed together, while in the English context, English-standard, English deviant-within and English deviant-between were analyzed together. Data is presented in this order within each time window.

2.3 Results

We conducted two experiments in which the frequency of the deviants differed (see methods). However, analysis of the two groups did not differ, and therefore, all reported results are for a total of 47 participants (26 participants in the original design group, and 21 from the control group).

2.3.1 Behavioral results

In the memory confusion paradigm (MCP), the types of errors participants made were collected and analyzed. Errors were differentiated as same-language errors or different-language errors. The logic behind the MCP is that if participants indeed categorize by language, they are more likely to confuse faces from the same-language than the different-language (i.e.- confuse the Spanish faces with other Spanish faces, than Spanish with English faces), making more same-language than different-language errors (this paradigm will be explained further in Chapter 3). While there are only three possibilities of making a same-language error (one of the faces is a correct answer), there are four possibilities to make a different-language error. To correct for this discrepancy, the between-language error was multiplied by 0.75.

Recognition Test 1

In the first recall test, participants made an average of 18.3 total errors (SD=2.8) out of 24 responses (75% error rate). As predicted, participants significantly made more same-language errors (10.1, SD= 2.8) than different-language errors (6.2, SD= 3.0, $t(46)=5.006$, $p<0.001$) (figure 2.2). This indicates that participants did indeed categorize faces according to the language they were accompanied with.

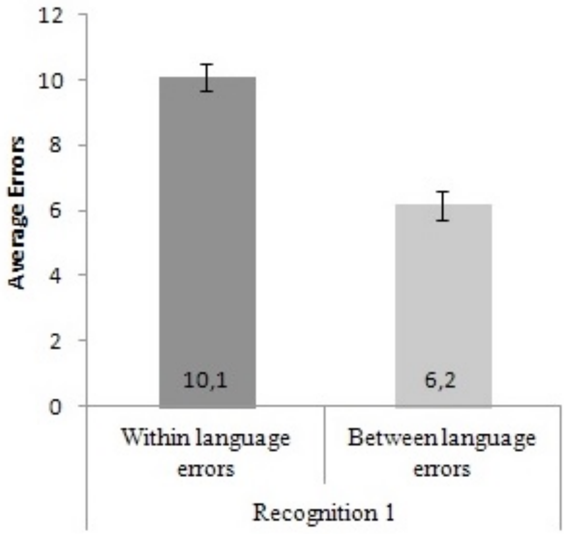


Figure 2.2-Within and between language errors for Recognition 1**.**

To see if there was an effect of language on responses, a repeated-measures ANOVA was conducted with two factors: language (Spanish and English) and response type (within-language error, between-language error and correct response).

No main effect of language was found (F(1,46)=1.231, *p*=0.273) signifying there were no differences in response to Spanish and English faces (see figure 2.3). A main effect of response type was found (F(2,92)=24.021, *p*<0.001) meaning that participants made much fewer accurate responses, and made many more within-language errors. A post-hoc, pairwise comparison showed that the accurate responses were significantly different from within-language errors (*p*<0.001), and that within and between-language errors were different (*p*<0.001) but that accurate and between-language errors were not different (*p*=0.628).

Finally, we found an interaction between language and response type (F(2,92)=3.945, *p*=0.023). A pairwise comparison of the two languages across response types showed that accuracy rates were higher for Spanish faces (3.3, SD=1.8) than for English faces (2.4, SD= 1.7, *p*=0.01). All other comparisons were not significant (*p*>0.137).

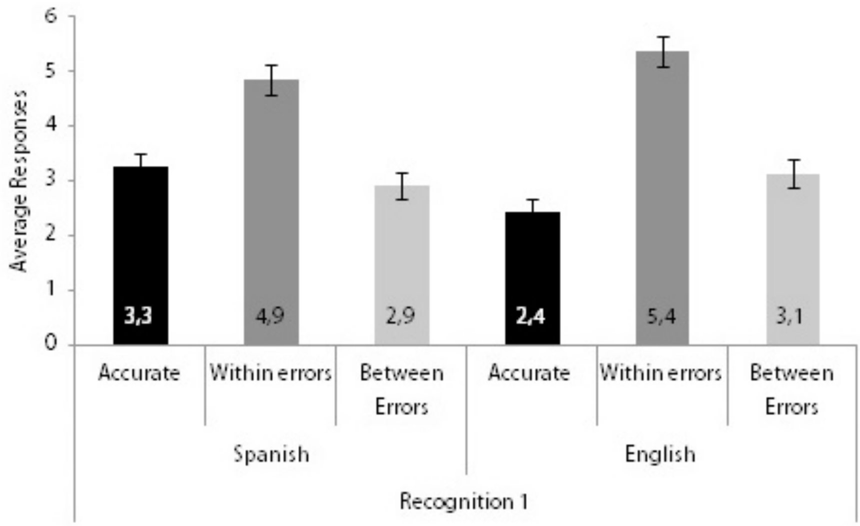


Figure 2.3-Average responses in Recognition 1 split by language (Spanish and English). Average responses in recognition 1 and 2. Between category errors are multiplied by 0.75 to correct for different base rates. Error bars denote standard errors.

Recognition Test 2

In Recognition 2, data from only 46 participants were collected due to technical malfunction of the program in one participant. Session 1 and Session 2 of the Recognition Tests were compared using a repeated-measures ANOVA with three levels: Response Type (accurate, error-within, error-between) Language (Spanish and English) and Session (Recognition 1 and Recognition 2). There was no difference between the two sessions (F(1,45)=0.031, *p*=0.826), meaning that participants performed exactly the same in both recognition tests. There was no effect of session and language (F(1,45)==0.14, *p*=0.707), or session and response type (F(2,90)=0.092, *p*=0.913), or a three way interaction between session, language and response type (F(2,90)=0.126, *p*=0.881).

2.3.2 Event-related potentials

Figure 2.5 shows the ERPs elicited by the three stimuli types (standards, within-language deviants and between-language deviants) across languages in three representative electrodes (Fz, Cz and Pz). Two time windows were selected for analysis based on visual inspection of the grand averages, and due to previous literature using the oddball paradigm, which were the 110-150ms and 180-310ms time windows, the early stages of face perception (vMMN) and post-perceptual stages (P3 component), respectively. The vMMN is located in posterior regions, and therefore, posterior electrodes were chosen for the analyses within the early time window. The 180-310ms time-window was analyzed considering different regions since there were no a priori hypotheses as to where the effect was localized. We were interested in the effect of condition, and therefore, we reported values when a main effect of condition or an interaction of condition with another factor was found.

We first analyzed the data as a language collapsed condition and later split each time window by the linguistic context in which faces were presented.

110-150ms average amplitude analysis

The average amplitude was analyzed in the 110-150ms time window. A repeated measures ANOVA with 2 factors: condition (standard, deviant-within and deviant-between), and electrode (CP5, CP1, CP2, CP6, P3, Pz, P4) were selected. There was no main effect of condition ($F(2,92)=2.125$, $p=0.125$).

In order to compare the linguistic contexts the faces were presented in, we did an additional analysis with an added factor of language context (Spanish standard context or English standard context). We found an effect nearing significant of linguistic context

(F(1,46)=3.993, *p*=0.052), but no interaction between linguistic context and condition (F(2,92)=1.32, *p*=0.272) (see figures 2.7 and 2.8).

110-150 peak amplitude analysis

A peak amplitude analysis was conducted within the 110-150ms time window. The same repeated-measures ANOVA was conducted for the peak analysis for this time window (3 conditions and 7 electrodes). Results showed a main effect of condition (F(2,92)=3.274, *p*=0.042), and pairwise comparisons [1] of the conditions showed that the deviant-between was significantly more negative than the standard (*p*=0.028), while other comparisons did not differ (*p*>0.46) (see figure 2.4 for a topographical map of this time window).

Another analysis comparing the two linguistic contexts was done for the peak amplitude in this time window (same analysis as done for average amplitude in the same time window). However, we did not find any significant results (*p*>0.137). Therefore, language contexts did not affect the peak amplitude within this early time window.

[1] All post-hoc pairwise comparisons were conducted within the ANOVA analyses, and multiple comparisons were corrected by Bonferroni methods. Therefore, only p-values are reported for pairwise comparisons.

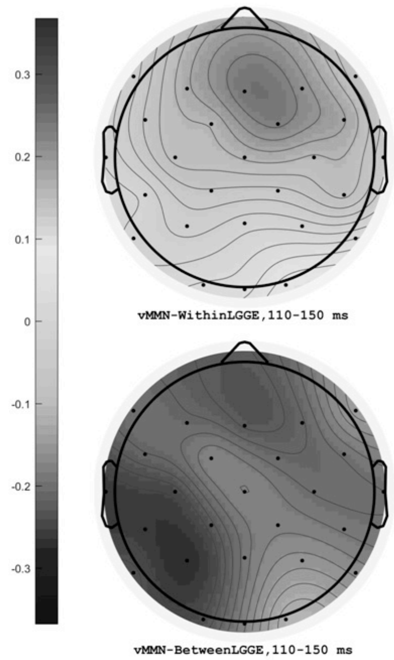


Figure 2.4- Topographical map of electrodes between 110-150ms. Effect shown here is the visual mismatch negativity (vMMN) effect in which the standard is subtracted from the deviant. The top map shows the vMMN for deviant-within language, and the lower map presents the vMMN of the between-language deviant. As can be seen, the effect is mainly in the posterior electrodes for the between-language deviant.

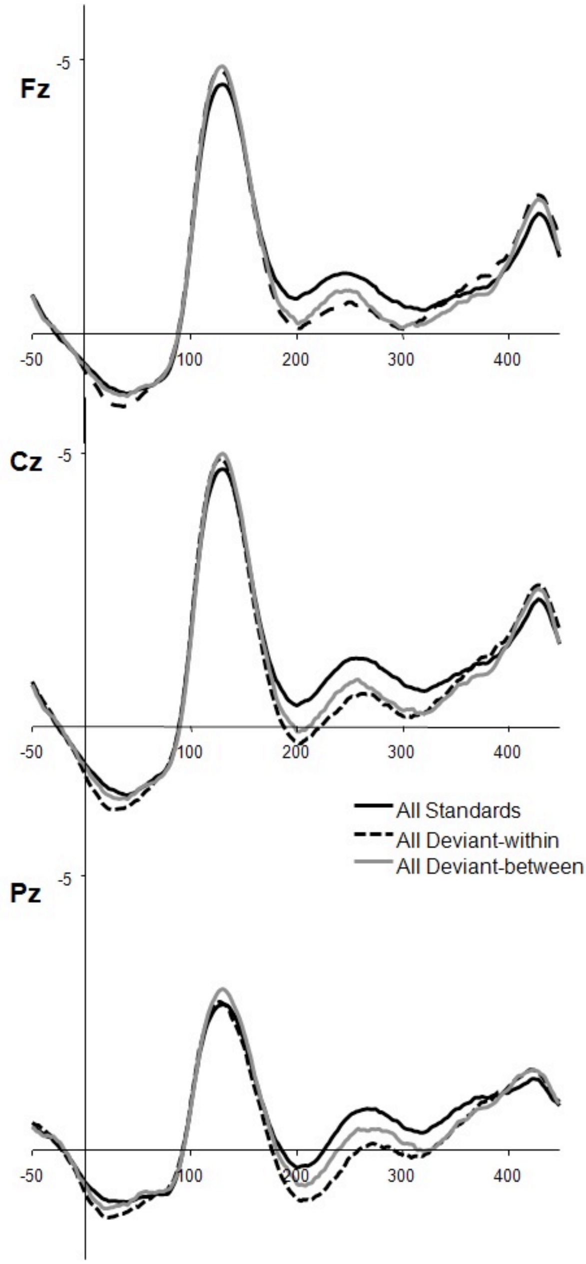


Figure 2.5- ERP waveforms for three electrodes Fz, Cz and Pz. ERP wave forms for all three conditions (standard, deviant-within and deviant-between) are shown from -100ms of time of onset until 450ms time post stimulus onset.

180-310ms amplitude analysis

The average amplitude within the time window 180-310ms was conducted. A repeated measures ANOVA with three conditions (standard, deviant-within, deviant-between), two regions (frontal and posterior) and six electrodes (fronto-central: F3, Fz, F4, FC1, FC2, Cz, posterior: P3, Pz, P4, O1, Oz, O2) was conducted. Results showed a main effect of condition ($F(2,92)=11.0$, $p<0.001$), and a pairwise comparison of the three conditions showed that the standard elicited a significantly smaller positivity compared to both the deviant-within ($p<0.001$) and deviant-between ($p=0.039$), but the two deviants were not significantly different from each other ($p=0.099$) (see figure 2.5). There was an interaction between condition and region ($F(2,92)=3.738$, $p=0.027$), and a pairwise comparison of the conditions in these two regions showed that in the frontal region, the standard condition was significantly different from the deviant within ($p<0.001$) and the deviant-between ($p=0.006$), but the two deviants did not differ from each other ($p=0.592$). In the posterior region, only the deviant-within was significantly different from both the deviant-between ($p=0.016$) and standard ($p=0.001$), while the standard and deviant-between were not significantly different ($p=0.549$).

This shows that the deviant-within showed the most robust effect across regions in this time window (see figure 2.6 for a topographical map of this time window showing that deviant-within had a larger effect).

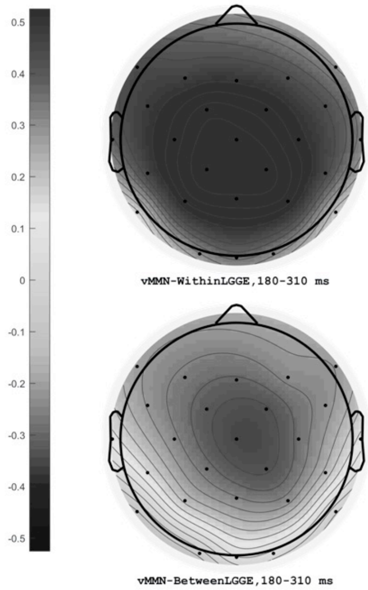


Figure 2.6-Topographical map of 180-310ms time window. The top map shows the deviant-within language minus the standard condition, and the bottom map shows the deviant-between minus the standard condition. These two maps show that the deviant-within had a larger and more regionally robust effect.

Finally, linguistic context was also compared within this time window, and therefore, the same repeated-measures ANOVA from this time window was conducted with an additional factor of context (Spanish or English standard contexts, see figure 2.7 and 2.8). However, we did not find a main effect of linguistic context ($F(1,46)=0.296$, $p=0.589$) or an interaction of the context with any other factors ($p>0.359$). Therefore, this time window was not affected by the language of the faces that were presented, but only by the deviancy from the category.

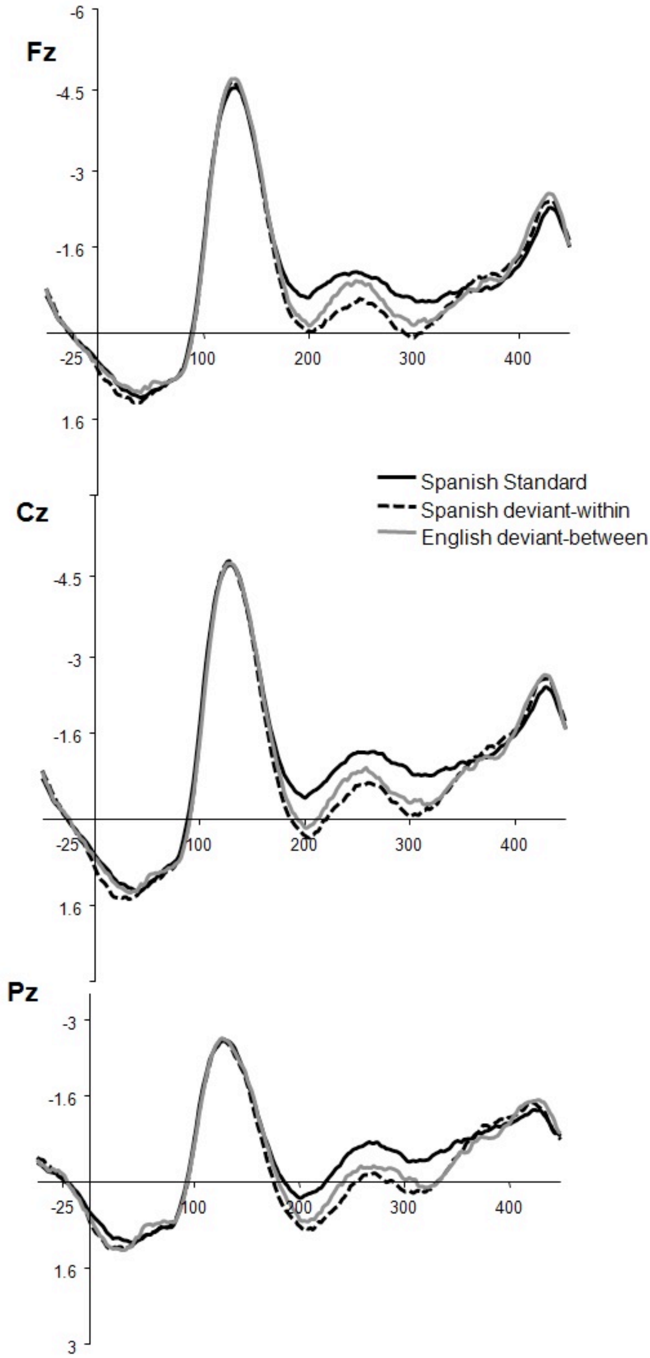


Figure 2.7 ERP graphs from Spanish context for three representative electrodes: Fz, Cz and Pz. Negativity is plotted upwards.

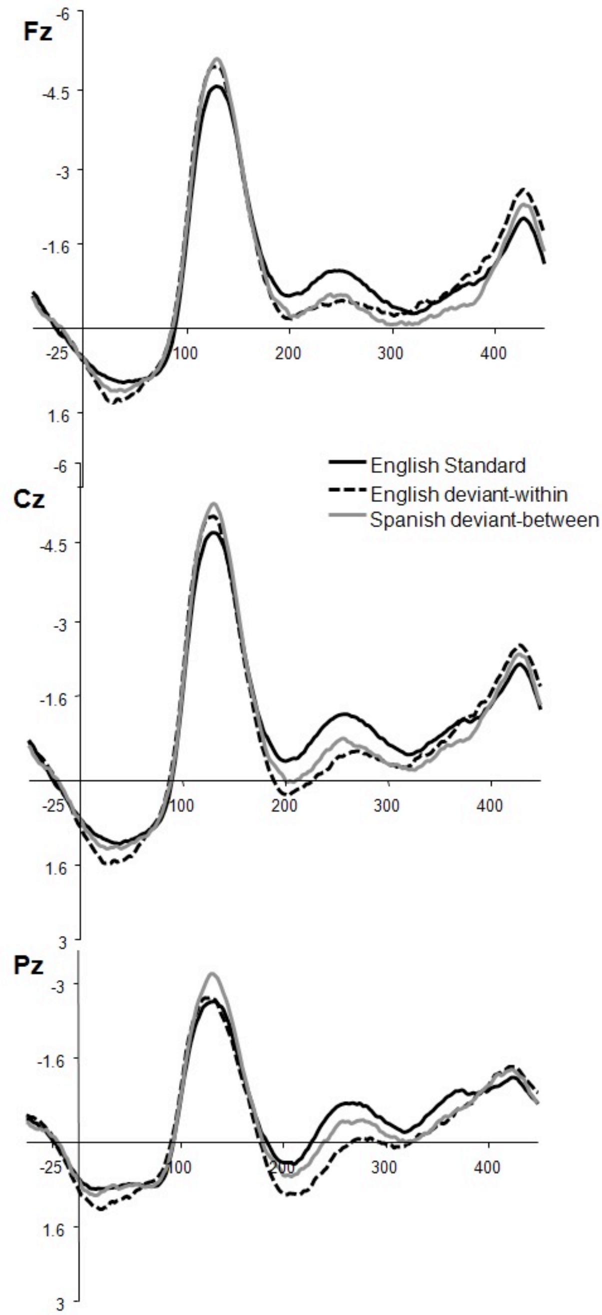


Figure 2.8 ERP from the English language context with three electrodes: Fz, Cz and Pz.

2.4 Discussion

2.4.1 Categorization of faces

In a series of behavioral and EEG experiments, we examined whether language is used as a social category to classify faces, and whether this language categorization affects low-level perception of faces. In the memory confusion paradigm (MCP), participants were more likely to confuse faces from within the same language group than from the other language category. That is, faces that were presented with Spanish (Spanish faces) were more likely to be confused with other Spanish faces than with English faces, and English faces were more likely to be confused with other English faces than Spanish faces. The MCP relies on these error types, and posits that if participants make more within than between-category errors, they have implicitly created categories. We showed, as many studies have previously done with accents (Pietraszewski & Schwartz, 2014b, 2014a; Rakić et al., 2011), that language was also used as a social category.

Interestingly, this categorization effect was present over an hour after its initial exposure to the language categories. The fact that we found the same effects in both recognition phases that were over an hour apart signifies that participants maintained the categories throughout the oddball paradigm. Therefore, we can be certain that the ERP results obtained in the oddball paradigm are due to the effects of categorization.

2.4.2 Effect of language categories on visual perception of faces

The MCP paradigm allowed us to conclude that faces were indeed categorized by language. Therefore, by using the oddball paradigm, we examined whether those newly created language categories affected visual perception of faces.

The early stages for face perception revealed a category effect in which deviancy from the *category* elicited a larger peak than standards or the within-category deviants. This signifies that participants detected a deviancy from the *category*, and not merely a deviancy in frequency of presentation (i.e. frequent versus infrequently seen faces)[2]. Previous studies on the effect of color categories on visual perception have also found category effects in the early time windows, in which the between-category deviant showed more negative waveforms compared to the within-category deviant (Holmes et al., 2009). This signifies that the language category of a face is detected within the early stages of face processing.

While statistically insignificant, visual inspection of the early time window seems to indicate that the Spanish deviant-between in the English context seems to have the most robust effect (see fig. 2.8 Pz electrodes). While this is speculative, perhaps specific language contexts affect early visual processing of faces in different ways. The categorization-individuation model (Hugenberg, Miller, & Claypool, 2007; Hugenberg et al., 2010; Sporer, 2001) posits that there are two ways of processing faces: categorization and individuation. Simply, categorization classifies faces (or other stimuli) into a group that shares the same dimension, while individuation allows the discrimination of exemplars within that category. Previous research has shown that own-group faces are

---

[2] In addition, this was true for both the original paradigm in which deviant-within was the same face presented 48 times, while the deviant-between was four faces presented 12 times each (for a total of 48 deviant-between presentation) and the control experiment in which both the deviant-within and deviant-between were one face from each language which was presented 48 times each.

individualized, whereas out-group faces are not (Bernstein et al., 2007). Therefore, the faces that appeared with Spanish may have been individualized as they were own-group faces, whereas out-group English faces were merely categorized, allowing the participants to detect a Spanish individual among English faces. In simple terms, Spanish faces that had become 'individuals' stood out in a group of homogenous English faces, but not vice versa. This may have ecological validity, since an out-group member within a majority in-group situation (such as foreigners within your native context) may not be relevant to detect, whereas having an in-group member within a majority out-group situation may be beneficial. Hence, within this time window, not only was deviancy from *category* detected, but also *which* category was being violated. With stronger statistical power, future studies may benefit from examining the linguistic context on early stages of face processing.

While we found a language category effect in the peak amplitude of the early time window, the effect was very small. This was possibly due to the convoluted design of our oddball paradigm procedure. First, unlike classical oddball paradigm studies where there is one standard stimuli repeated frequently with one (or two) other deviant presentation, our design had three rotating face standards from a category and several deviant faces. In addition, our stimuli were of faces, which is arguably a more complex visual stimuli than, say, colors. These complex visual stimuli may have affected the size of the early ERP component. Secondly, all previous studies have used stimuli in which the category was visible within the stimuli (for a review see Czigler, 2014). For example, in the case of gender and emotion, the faces themselves reveal category cues in a bottom-up manner, whereas in our study, the faces themselves did not carry any visible language category information. This was done to be certain that category effects were truly a top-down effect, but it may have decreased the sensitivity of the effect. Thirdly, these language categories were implicitly created on-site at the laboratory, in an encoding phase that only lasted approximately two minutes. There were no explicit instructions to learn which face belonged to which language category, and there were no explicit training sessions, unlike other studies which trained participants to learn categories (Clifford et al., 2012). For these reasons, it is not surprising that the effect in the early visual component is small.

In sum, top-down effects of language categories were found in the early stages in which category deviants elicited the most negative peak amplitude.

Language categories were also reflected in the post-perceptual time window, but in a different manner from the early stages of face processing. Firstly, we found a general deviancy effect, in that both deviants were significantly different from the standard. Secondly, we found a larger effect in *within* category deviant, reflecting a detection of language category. Therefore, this time window reflects both the detection of deviancy from frequently seen faces as well as a deviation from the language category.

While we found deviant effects in the post-perceptual time window, the direction of magnitude among the two deviant types do not go in line with previous studies (Clifford et al., 2012; Yu et al., 2017). Clifford and colleagues (2012) trained one group of participants with new categories for green, while another group did not. They then examined whether the newly learned color categories elicited a within- and between-category deviant effect during an oddball paradigm. Category effects of newly learned color categories were only found in the trained group, and this was only seen in what they called the post-perceptual stages of processing (P3 time window at 350-600ms). Their results show that the *between*-category deviant elicited a more positive wave than the deviant-within. Due to their findings, they argue that newly learned color categories affect cognitive mechanisms, but that they are independent from early perceptual processes. While our deviant faces also elicited a positivity in a similar time window, the pattern of our results differ from those in Clifford (et al., 2012). We found that the faces from *within* category deviants elicited a greater positivity in the late time window compared to the *between* category deviant (180-310ms).

The directional difference between our study and previous studies may be due to the convoluted design of our oddball paradigm. Studies testing newly learned categories used one standard deviant with either one deviant at a time (Yu et al., 2017) or two deviants (Clifford et al., 2012). For example, Yu and colleagues (2017), tested newly learned categories of faces, but only used one face as a standard interrupted by either a deviant within or deviant between face (i.e. only two faces seen per block). This design may maximize the effect of deviancy from the standard, but because the standard identity is always the same (not counterbalanced), the effect between the two deviants may also reflect a detection of individual identities. In contrast, we used three alternating standards interrupted by both within and between deviant types for a total of five of six different faces within a block[3]. All faces were presented as all conditions for each participant. By doing so, we remove the effect of deviancy created by individual faces, and insured that effects were purely due to the category of the faces. Perhaps this difference affected the later time window as an increased cognitive demand, or complex visual stimuli. Since category effects were detected in the early time window, perhaps the later time window reflects extra processing of faces whose categories were not detected. There may also be a special effect of *language* categories on faces when compared to other category types. These explanations are speculative, and therefore future studies should attempt to disentangling these effects and examine how categories (language and otherwise) affect the visual perception of faces.

---

[3] Three deviants were presented in the original design, and two deviants were presented in the second design. This difference is explained in the methods.

In summary, the later time window shows both a detection of deviancy in frequency, as well as a category deviancy effect. Therefore, our results suggest that language category also effects affect post-perceptual time windows in face processing.

2.4.3 Summary and future directions

Our current study shows, behaviorally, that faces are categorized by the language they accompany. We thereby add language as another dimension of social categories that have been tested including race, gender, age, university-affiliation (Bernstein et al., 2007), socio-economic class (Shriver et al., 2008) accents (Pietraszewski & Schwartz, 2014b) or coalitional groups (Pietraszewski & Schwartz, 2014a).

Secondly, our results found language category effects in early ERP components. Therefore, our results show that social categorization affects both pre- and post-perceptual stages of face processing. Furthermore, we add more support to the increasing body of literature claiming that categories affect early visual perception in a top-down manner. Previous studies on the effect of newly learned categories on visual perception have mostly used color categories (Clifford et al., 2012). Our study extends these findings by demonstrating that categories of complex visual stimuli, such as faces, can also affect low-level, early stages of visual perception.

In this chapter of the dissertation, we have established that language is used in social categorization of faces, and that this categorization affects the visual perception of these faces in a top-down manner. In the following chapter, we examined the interaction of language categorization with face categorization, as defined by race. That is, now that we have established that languages can be used to categorize faces (of the same race), do we do so for other race faces? Additionally, when two languages and two races are crossed to create multiple-group membership individuals (e.g. Caucasians who speak English (L2) versus Black people who speak Spanish

(L1) are language or race categories more robust? In the next chapter, we used the memory confusion paradigm to test categorization effects in single as well as multi-dimensional condition.

Chapter 3

The interaction of language and race in social categorization.

3.1 Introduction

Social categorization is an automatic and instantaneous phenomenon that happens when we meet new people. Among the multitude of ways in which we can categorize someone, decades of research has shown that there are the 'big three' social categories that are automatic, instantaneous and unavoidable: race, gender and age (Allport, 1954; Hills, 2012; Meissner & Brigham, 2001; Palmer et al., 2013; Rhodes & Anastasi, 2012; Wright & Sladden, 2003). On the other hand, the Ethnolinguistic Identity Theory (ELIT) (GILES & Johnson, 1987) states that language is one of the most, if not *the* most important aspect of social categorization of self and others. In the work of Stevenage, Clarke, and McNeill (2012), it has been found that participants were able to recognize voices from their own-accent (English vs. Scottish) better than voices from the other-accent. This phenomenon was named the Other-Accent Effect (OAE), due to its similarity with the Other-Race Effect (ORE) (Allport, 1954) in which recognition for the group with the 'other' characteristic is decreased. Recent research has argued that the "big three" is incomplete, and that language could be considered the fourth dimension in social categorization (Kinzler et al., 2010).

These social categories are used to classify people into in/out-group members spontaneously, where in-group members are better recognized than out-group members (eg. Meissner & Brigham, 2001; Stevenage et al., 2012). In experimental settings, group memberships can easily be controlled for, but in reality, people belong to multiple social categories, in which they may be in-group in one dimension, but out-group in another. For example, someone from your own-race could be from another age-group, consequently inhabiting multiple categories at once. These multiple-group members complicate the patterns of social categorization, especially when considering how viewers project these multi-group individuals into in/out group memberships relative to themselves. Yet, within the reality of modern society, it has become more common to engage in daily interactions with others from diverse racial backgrounds in workplaces or in classrooms. On the other hand, encounters with individuals of the same-race who inhabit other social categories like politics, nationalities, religion, or sports-teams is also common. With increasing scales of globalization, it can be expected that these myriad types of group interactions will be on the rise, and necessitate the examination of social categorization reflecting the more modern realities of daily life.

One way to measure social categorization is by examining recognition memory of faces. Several studies have explored the interaction of race with other social categories such as university affiliation or their socio-economic background (Hehman et al., 2011; Shriver et al., 2008). Shriver (et al., 2008) found that among White participants, recognition memory for White faces accompanying higher socio-economic backgrounds was improved compared to White faces accompanying low socio-economic backgrounds. However, higher socio-economic backgrounds did not improve recognition for Black faces. These results show a *social-exclusion pattern* in which double in-group members (ie- rich White faces) are remembered better over all other groups (partial in-group members and double out-group members; ie- low socio-economic

White faces and Black faces regardless of their socio-economic status). The authors suggest that impoverished White faces became categorized as out-group members, and therefore recognition decreased, but for Black faces which were already categorized as racial out-group members, socio-economic context did not affect this membership and therefore its recognition. This finding suggests the strength of race categorization over socio-economic membership.

On the other hand, Hehman and colleagues (2011) found that when race was crossed with university-affiliation, recognition for own-race faces was not affected, but own-university affiliation helped recognition for other-race faces, showing a *partial in-group pattern* of multiple categorization, meaning that in some contexts, recognition for otherwise out-group members (Black faces) can be improved by becoming in-group members in another dimension (university affiliation). In all these studies however, recognition for own-race faces were always better than other-race faces regardless of other social cues. In sum, while past research has shown that other social categories interact with race to create different patterns of social categorization and memory for multiple-group members, these studies also indicate the robustness of race on recognition for people with multi-group memberships.

Other studies more aligned with the ELIT have examined the effect of accent when categorizing faces (Pietraszewski & Schwartz, 2014a, 2014b; Rakić et al., 2011), and have found that accents are a robust dimension of social categorization. In these studies, authors used the Who Said What (WSW) paradigm (also known as the Memory Confusion Paradigm, Klauer & Wegener, 1998; Taylor, Fiske, Etcoff, & Ruderman, 1978), a popular psychological tool used to measure implicitly created social categories. In this paradigm, several faces are shown from two implicit groups, for example American vs. British accented speakers (Pietraszewski & Schwartz, 2014b, 2014a). All studies consistently found that participants created category boundaries according to the accent in which faces were presented, even in the absence of explicit instructions. Furthermore, Pietraszewski and colleagues (Pietraszewski & Schwartz, 2014a) have shown that when accents (American vs. British) were crossed with coalitional information (charity groups) to create multi-group members (i.e.- American speaker from charity group A, American speaker from charity group B, British speaker from charity group A and British speaker from charity group B) participants were more likely to confuse speakers among accent groups rather than charity groups, suggesting the robustness of accent as a social category, even when given other social membership information. However, when race (White and Black faces) was crossed with coalitional information (charity group membership) to create similar multi-group members, racial categorization strength decreased. Based on these results, it was concluded that accent is a robust dimension of social categorization with evolutionary origins, whereas race is an evolutionary byproduct of detecting coalitional groups.

What is notable, however, are the lack of studies directly comparing the effect of race and language and their interaction in creating social categories: that is, when race and language are simultaneously available cues. Simply put, how do individuals categorize others from another race who speak their language, and others from their own race who speak a foreign language? As mentioned previously, modern society makes it increasingly likely to meet people from the same racial background who speak a different language, while it is also possible to meet people from other races who speak our native language. The only study which directly compared these two groups (own-race, other-language faces vs. other-race, own-language faces) have been conducted on children (Kinzler et al., 2009). When five-year olds were given the choice between a playmate from their own race with a foreign-accent, or a playmate from another race with a native-accent, the children chose the latter. The authors concluded that languages and accents are important features in guiding social preferences in the case of a child's choice for playmates. Note however, that children were not given the option to play with same-race, same-accent playmate (double in-group) or other-race, other-accent playmates (double out-group).

No such study directly crossing language and race has been conducted on adults. The only study available in adults compared accents with 'ethnic looks' (Rakić et al., 2011). Rakić et al., (2011) crossed accents (native and foreign) with a person's 'ethnic look' (typical German looks vs. typical Italian looks). They first tested whether participants categorized faces to the same degree based on typical ethnic looks (typical German or typical Italian faces) and accents (native- or foreign-accented German) and found that both looks and accents trigger social categorization. In a second experiment, the authors crossed the looks and accents to create four groups of faces (typical German looking face with native-accented German, typical German looking face with an Italian accent, typical Italian face with a native-accented German, and typical Italian face with an Italian accent) to test the influence of looks and accents on ethnic categorization when combined. Results showed that participants had better memory for accents than for looks, allowing authors to conclude that accents provide more relevant information in categorization when compared to 'ethnic looks', supporting the ELIT. This study is often cited as supporting the idea that visual cues are of minor relevance in social categorization when compared to linguistic cues. However, both faces used in this study were from the same Caucasian race (German and Italian). This may not be sufficient to conclude that accent or language information surpasses *race,* a much more visually salient cue than 'ethnic cues'*,* in social categorization.

Therefore, the aim of this current study is to examine the categorization strength of language and race when these categories are crossed. To do this, we used the Memory Confusion Paradigm (otherwise known as the 'Who Said What' paradigm; K. C. Klauer & Wegener, 1998; Taylor et al., 1978). In this paradigm, participants were shown a set of faces with different characteristics, each accompanied by a sentence. In the one-dimensional paradigm, half of the faces shared one characteristic, and the other half shared another characteristic (e.g. race; White men and Black men). In the two-dimensional paradigm, characteristics were crossed to give four groups. In our study, for example, we used: 1) White men speaking Spanish, 2) White men speaking English 3) Black men speaking Spanish and 4) Black men speaking English. In a second phase, participants underwent a recognition test where they were presented with all the faces and the statements they previously heard, and participants answered 'who said what'. Participants report this task as being difficult, and perceive their responses as random guesses. However, when errors are analyzed, the errors tend to fall within a certain pattern. When an error is made, it is more likely to be a face that shared the same characteristic as the correct answer, than a face that does not share that characteristic. That is, if categories were created, members who share the same characteristics are more confused amongst each other. In this way, it is possible to measure implicitly created social categories. By using a two-dimensional WSW paradigm, it is also possible to look at the strength of a characteristic when it is crossed with other cues. By comparing categorization strength in a one- and two-dimensional WSW paradigm, we can examine the robustness and retention of social categories.

The main objective of our study was to examine the interaction of language and race in creating social categories. To this end, we also examined the effect of language categories in both an own- and other-race condition (White and Black races respective), and the effect of race categories in a Spanish (L1) and English (L2) context. Finally, we also examined in/out-group effects across conditions. Previous studies have shown that participants indeed do categorize people by accents or race, but these studies have only examined accent categorization among own-race faces, and race categorization in an own-accent (own-language) context. Therefore, we added two extra conditions where these were controlled: language categorization in an other-race condition and race categorization in an L2 condition. To examine all of these questions, we conducted five experiments exploring the interaction of social categorization of language (Spanish and English) and race (White and Black). In Experiments 1 and 2, we examined the categorization strength of language in either an all-White faces condition (Experiment 1) or an all-Black faces condition (Experiment 2). This allowed us to test the categorization strength of language not only in an own-race situation (Experiment 1) but also in an other-race situation (Experiment 2) to see whether language categorization strength remains the same across different race faces. In Experiments 3 and 4, we investigated the categorization size of race in an all-Spanish speaking or all-English speaking condition, respectively. In these experiments, we explored the categorization strength of race in a Spanish (participants' native language) condition and whether it remains robust in an English (participants' foreign language) condition. These four experiments acted as controls and baseline conditions for the two-dimensional experiment. Finally, in Experiment 5, we crossed the two characteristics to examine which dimension remained robust when both cues were available for categorization. Finally, we questioned whether in/out-groups across language and race conditions affected face memory, especially when comparing double in-group, double out-group, and partial in-group conditions.

3.2 Methods and Results

3.2.1 Methods for all experiments

In order examine the interaction of race and language when creating social categories, the Memory Confusion Paradigm (MCP) was used following the methods of Pietraszewski and Schwartz (2014).

3.2.2 Stimuli

Grey scale photos of eight White males and eight Black males with neutral expressions were selected from free, electronic databases, which were downloaded from the web. They did not have any easily distinguishable facial characteristic (such as a mole) on their face. They all had dark hair and eyes, so that they would be convincing as either a Spanish or English-speaking person. Adobe Photoshop (version CS3 for Windows) was used to isolate their face and hair from the background and the photograph was digitally edited in order to make their hairstyles as similar as possible.

Twenty-four neutral, non-autobiographical sentences were created in Spanish, as well as its English translation of each (e.g.: El libro tiene cien páginas. [English: The book has a hundred pages.] For the full list see the Appendix). On average, Spanish phrases were 4.9 words in length, while English phrases were 5.1 words in length. No statistical difference was found between sentence lengths of the two languages ($t(23)= -1.072$, $p=0.3$). Both Spanish and English phrases ranged from four to seven words per phrase. Phrases were recorded and edited using Audacity (v 2.0.3) from eight native Spanish speakers for the Spanish sentences, and eight native English speakers for the English sentences.

3.2.3 Participants

A total of 159 participants from the database of the Pompeu Fabra University (Barcelona, Spain), mean age 21.4 (SD= 2.5) who were Spanish dominant speakers participated in this study. All experiments had 32 participants each, except for experiment 3 which had 31. Sex of participants was collected in Experiments 1, 2 and 3, which had 19, 3 and 11 males respectively. However, sex was not collected for Experiments 4 and 5. Participants from Experiment 1 are the same participants from the original oddball paradigm experiment in Chapter 2, and data from all 32 participants were analyzed.

3.2.4 Procedure

Participants were seated in front of a computer screen in a soundproof room. The experiment was presented to the participants using the software E-prime 2.0 (Schneider et al., 2002). The sentences were presented auditorily through headphones.

The memory confusion paradigm consists of three parts. The first is the encoding phase, the second is the distraction phase and finally, the recognition phase. During the encoding phase, individual photos were presented on the screen while the phrases were presented auditorily through headphones. Participants were told to form impressions of the speakers while they watched and listened, as they would be asked questions later regarding this portion of the experiment.

Each of the eight faces appeared three times during the encoding phase, for a total of 24 presentations. The three presentations of each face had a different sentence, but voices remained the same. Photos and audio were presented simultaneously, and the photo remained on screen for a total of 4010ms. This time was selected to allow the face to appear an extra 2000ms after the longest sentence, which was 2010 ms in length. A blank grey slide appeared for 200ms between each photograph (for the procedure see figure 3.1).

After the encoding phase, the participants played one minute of Tetris as a distraction from the task.

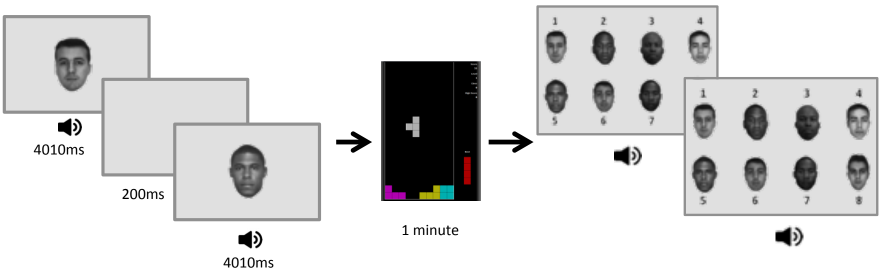


Figure 3.1 The procedure of the Memory Confusion Paradigm (MCP). This diagram shows the MCP in a mixed race condition (ie- experiment 3 and 4). The very left is the encoding phase, where faces were presented with audio. Then participants played one minute of Tetris as a distractor phase, and then finally did the recognition phase.

During the Recognition phase, all eight photos were presented in one screen, numbered 1 to 8. At the same time, the same sentences were presented auditorily. The participant had to decide which of the eight faces accompanied that sentence in the Encoding phase by pressing the corresponding number on the keyboard (i.e.- who said that sentence). The eight faces remained on screen until the participant responded. After the response, there was a blank screen for 1000ms. This was done until all 24 sentences that the participants heard in the encoding phase were presented. Sixteen lists were created to counterbalance the face, sentence and language. Therefore, all faces accompanied every sentence in both languages across participants.

3.2.5 Analyses – correcting for base-rate error probabilities

In every recognition phase, we collected accurate and incorrect responses for each face, and each error was categorized as a within- or between-category error. In experiments 1 and 2, errors were categorized as within-language or between-language errors, and in experiments 3 and 4, errors were either within-race or between-race errors. In these one dimensional MCP experiments, the between-category error is multiplied by 0.75 to correct for different base-rate probabilities (i.e.- because there is one correct speaker, there are only 3 within-category error possibilities while there are 4 between-category error possibilities). Categorization is calculated by subtracting the between-category errors from the within-category errors. Therefore, for example, categorization by race is calculated as: (C_r) = (within-race error) - (between-language error x 0.75).

In Experiment 5 where there were two dimensions, a similar base rate correction was made. However, since there were four possible error outcomes (within-language/within-race, within-language/ between-race, between-language/within-race and between-language/between-race), a within-category error for language was calculated as [within-language/within-race + within-language/between-race], while between-category error for language was calculated as [between-language/within-race + between-language/between-race x 0.75] (Bor, 2018). The same was done for race.

Effect sizes were calculated by comparing the within- and between-category errors using this formula:

$$r_{ES} = \sqrt{\frac{t^2}{t^2 + (N-1)}}$$

Cohen (1988), proposes that effect size (r_{ES}) interpretations are as follows: 0.1 = small effect, 0.3 = medium effect and 0.5 = large effect. Having higher effect sizes means participants made more within- than between-category errors, signifying the strength of categorization.

3.2.6 Experiment 1- All White faces with English and Spanish phrases

Methods

In Experiment 1, all the faces were of White men. Half the phrases were Spanish phrases, and the other half were English phrases.

Four of the photos were accompanied with Spanish (L1) phrases, and the other four were accompanied with English (L2) phrases. The first two photos were always accompanied with English sentences and the third and fourth were always Spanish sentences. From then on, the presentation of speakers was randomized, with the constraint that each speaker was presented within presentation 1 to 8, and then again within 9 to 16, and finally between 17 to 24.

Results

In Experiment 1, by only having White faces coupled with either Spanish or English sentences, we examined whether languages were used to categorize faces. Participants made an average of 16.33 errors (SD=1.8). Of these errors, there were significantly more within-language errors (10.28 ± SD=2.81) than between-language errors (6.05 ± SD= 2.76, $t(31)= 4.54$, $p< 0.001$, $r= 0.63$), indicating that participants indeed did categorize faces by their language (see figure 3.2).

3.2.7 Experiment 2- All Black faces with English and Spanish phrases.

Methods

Experiment 2 was identical to Experiment 1 except for the faces. In this experiment, all eight faces were of Black male faces.

Results

Participants made an average of 15.99 errors (SD= 2.92), of which more errors were within the language category (10.16 ± SD=3.75) than between the language category (5.84 ± SD=3.37, t(31)= 3.76, *p*=0.001, r=0.56). Experiments 1 and 2 indicate that participants indeed did categorize faces by their language, regardless of whether the faces were from their own or other races (see figure 3.2).

3.2.8 Experiment 3- All Spanish phrases with White and Black faces.

Methods

In Experiment 3 all the sentences were in Spanish. Four White and four Black male faces were used. In the encoding phase, the first two faces were always that of Black men, and the third and fourth photos were of White men. Otherwise, ordering was the same as Experiment 1.

Results

In Experiment 3, we tested the effect of categorization by race by having two race faces (White and Black) accompanying only Spanish sentences. In this experiment, participants made an average of 14.8 errors (SD=3.16), where more errors were within-race errors (9.03 ± SD=2.63) than between-race errors (5.81 ± SD=1.85; t(30)=

5.50, $p < 0.001$, r=0.71). Therefore, participants did categorize these faces by race in a Spanish language context.

3.2.9 Experiment 4- All English phrases with White and Black faces.

Methods

Experiment 4 is identical to Experiment 3, except that all sentences were English sentences instead of Spanish. These English sentences were translations of the Spanish sentences in order to control for sentence content, and had been used in the other mixed language conditions (Experiments 1 and 2).

Results

Experiment 4 measured race categorization in an English language condition. Participants made an average of 17 (SD=2.48) errors. Again, participants made more within-race errors (9.5 ± SD=2.72) than between-race errors (7.5 ± SD=2.08, $t(31) = 2.72$, $p = 0.011$, r=0.44). Results from Experiment 3 and 4 indicate that participants indeed categorized faces by race in both a Spanish (L1) and English (L2) context.

3.2.10 Experiment 5- Both race faces and both languages.

Methods

Finally, in Experiment 5, we had a 2x2 condition with both races and both languages. Therefore, there were two White faces accompanied by Spanish sentences, two White faces with English sentences, two Black faces accompanied by Spanish sentences and two Black faces with accompanying English sentences. In the encoding phase, the first two faces were always Black faces, and third and fourth faces were always White faces. From thereon out, race order was randomized. Language order was randomized from the beginning. In the recognition phase, order of presentation was randomized. To calculate error numbers, we calculated within language errors as {within-language/within-race errors + within-language/between-race} and between language errors as {(between-language/within-race + between-language/between-race) x 0.75} as suggested by Bor (2018) when calculating error rates in a multidimensional MCP experiment. Calculations for race errors were calculated in the following manner: within race as {within-language/within-race + between-language/within-race} and between race errors were calculated as {(within-language/between-race + between-language/between-race) x 0.75}.

Results

Unlike Experiments 1 to 4, it is impossible to calculate corrected error rates without splitting them either by language errors or race errors. When looking at error types by language, our results show that participants significantly made more within-language errors (9.94 ± SD=2.34) than between-language errors (6.91 ± SD=2.5; $t(31)= 4.08$, $p< 0.00$, $r=0.59$) and when looking at race, participants also made significantly more within-race errors (9.09 ± SD=1.65) than between-race errors (7.55 ± SD=2.05; $t(31)= 3.09$, $p=0.004$, $r=0.49$; see figure 3.2).

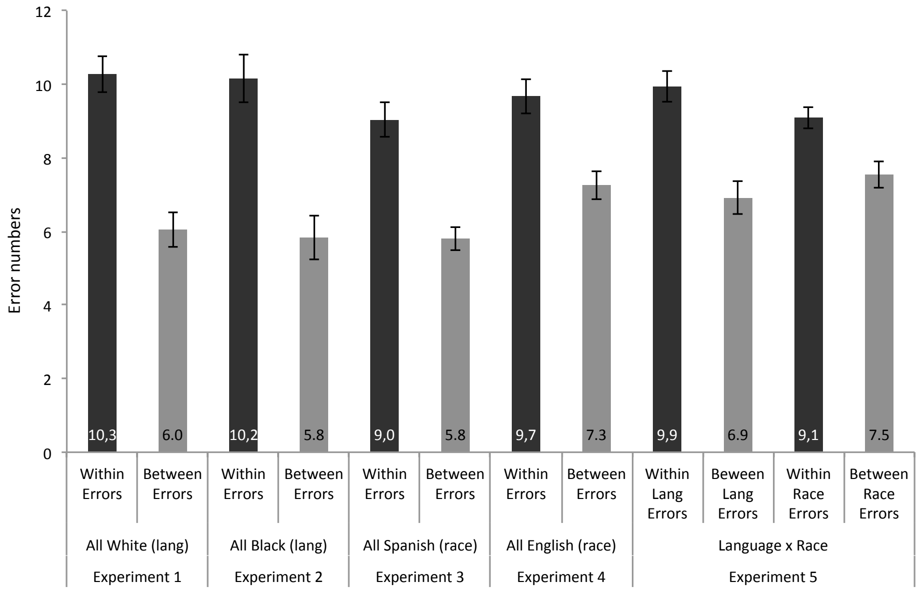


Figure 3.2- Average error responses (within and between category errors) for Experiments 1 through 4. The between category errors have been multiplied by 0.75 to correct for different base rates for Experiments 1 through 4, and the errors were split by language or race type in Experiment 5, and similar calculations were made. Error bars denote standard errors.

3.2.11 Comparing all Experiments

Categorization Strength

The categorization effect, as measured by effect size (r) indicates the strength of categorization of each characteristic, and allows us to compare categorization strength across single or multi-dimensional paradigms. The effect size is calculated by comparing within and between-category errors. Our results showed that the effect size (r) was large in all experiments (see figure 3.3), but especially for racial categorization in the all-Spanish condition (Experiment 3, r = 0.71). However, the lowest categorization effect was also for racial categorization in the English context (Experiment 4, r= 0.44). In the two dimensional experiment (Experiment 5), race categorization effect size was 0.49. This means that participants used race as the strongest cue for categorization in a Spanish context (Experiment 3), while racial categorization was relatively weaker in an English context (Experiment 4), and when two races and two languages are available categories, race becomes a weaker cue, comparable to that of an all-English context. Language categorization effect size was strongest in an all White condition (Experiment 1) at 0.63, then 0.56 in an all Black condition (Experiment 2) and 0.59 in the all-mixed condition (Experiment 5). This means that language categorization strength remains relatively stable across different racial and multi-group conditions.

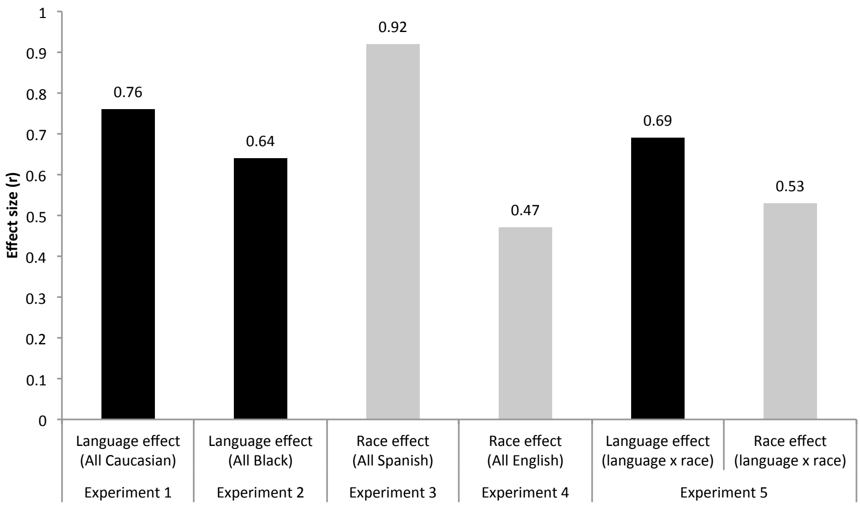


Figure 3.3- Effect sizes (r) for all five experiments.

3.2.12 In/Out-Group Effects

Normally in MCP experiments, accurate responses are not analyzed and in/out-groups are not explored. However, as an exploratory measure, we examined whether participants responded more *accurately* to Spanish vs. English faces in the all-White and all-Black conditions (Experiments 1 and 2 respectively) and White vs. Black faces in the all-Spanish and all-English conditions (Experiments 3 and 4) to test for in/out-group effects. To do so, paired t-tests were done comparing accuracy for White Spanish faces and White English faces in Experiment 1, which showed that participants moderately responded more accurately to White Spanish faces (3.25 ± SD= 1.76) than White English faces (2.51 ± SD =1.62; t(31)= 1.89, *p*=0.068). Accurate responses for Black Spanish (3.09 ± SD=2.35) and Black English (2.97 ± SD=1.93) faces in Experiment 2 did not show a significant difference (t(31)=0.273, *p*=0.786). These two results indicate that language differences did not have very strong in/out-group effects (see figure 3.4 A).

Accurate responses for White Spanish and Black Spanish faces in the all-Spanish condition (Experiment 3) were also compared in a paired t-test, showing that participants responded significantly more accurately to White Spanish faces (4.23 ± SD =2.36) than to Black Spanish faces (3.0 ± SD= 1.91; t(30)=2.791, *p*=0.009). Finally, accurate responses for White English and Black English faces in Experiment 4 were compared, but there were no significant differences (t(31)=-0.626, *p*=0.536). Together, this means that in/out-group effects were observed for race differences, but only in a Spanish context.

Finally, in Experiment 5, faces were split into four possible groups (White Spanish, White English, Black Spanish and Black English), and a repeated-measures ANOVA was conducted on accurate responses (figure 3.4 B) with 2 factors: language (Spanish and English) and race (White and Black). Note that possible accurate responses for each of these faces are half of that of Experiments 1 through 4 (i.e. - 6 possible accurate responses for each of the four face types in Experiment 5, but 12 possible accurate responses for two face types in Experiments 1 through 4). Results show that there were no main effects of race (F(1,31)=0.529, *p*=0.473) or language (F(1,31)=2.858, *p*=0.101), but an interaction of race*language was found (F(1,31)=10.343, *p*=0.003). This means that patterns of face memory were different across these four groups. As can be easily seen in figure 3.4 B, accuracy scores were highest for Black Spanish faces (1.6 ± SD=0.2) followed by White Spanish faces (1.4

± SD= 0.2), then White English (1.1 ± SD=0.2) and Black English (0.7 ± SD=0.2). These results are surprising in that the double out-group face (Black Spanish) had the highest accuracy score. However, it must be noted that while results are significantly different, ultimately, the accuracy scores are all extremely low, and difference between the highest and lowest accuracy scores are less than one response.

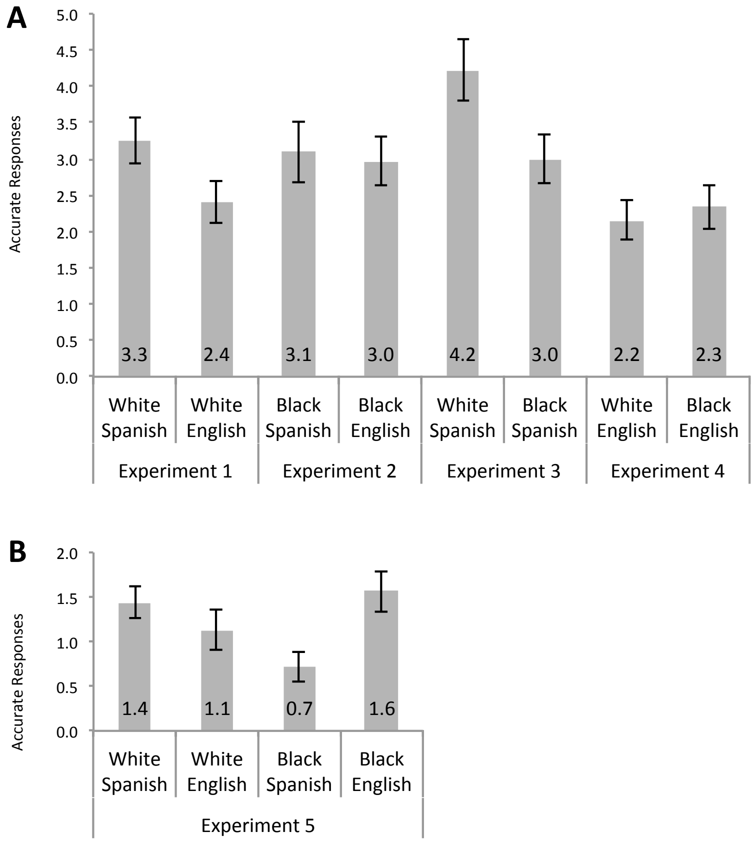


Figure 3.4 Average accurate responses across A) Experiments 1 through 4 and B) Experiment 5. The experiments are split because Experiment 5 (mixed condition) has half the number of possible accurate responses compared to Experiments 1 to 4. Error bars denote standard errors.

In order to compare the different experiments, a one-way ANOVA with Experiment as the between-subjects factor was conducted. The three possible responses (accurate, within-error and between-error) were individually examined in experiments 1, 2 and 5. Results showed that there was no difference among experiments in any of the response types (Accurate responses: $F(2,93)=1.445$, $p=0.241$; within-language errors: $F(2,93)=0.106$, $p=0.9$, between-language errors: $F(2,93)=1.244$, $p=0.293$).

Experiments 3, 4 and 5 were also compared in a one-way ANOVA with Experiment as the between-subjects factor with accurate responses, within-race errors and between-race errors as the dependent variables. Results showed that there was a significant difference in accurate responses ($F(2,92) =7.173$, $p=0.001$) and between-race error responses ($F(2,92)=7,729$, $p=0.001$), while there was no effect of within-race errors ($F2,93)=0.362$, $p=0.697$). A post-hoc Tukey HSD analysis showed that accurate and between-race errors for the all-Spanish condition (Experiment 3) were significantly different from the all-English (Experiment 4) and mixed condition experiments (Experiment 5) at the alpha level 0.05. All other comparisons were not significantly different. This means that the all-Spanish condition (Experiment 3) had the highest accuracy rate and lowest between-race errors across these three experiments.

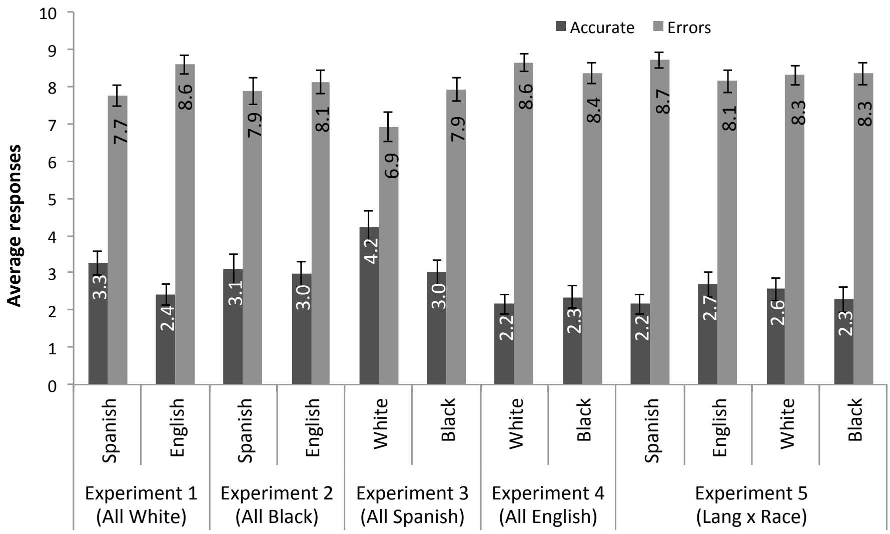


Figure 3.5- Average accurate and error responses across five experiments. Accuracy scores and incorrect responses are split by race and language to examine whether in/out-group effects were observable.

Finally, to look at whether responses to multi-group faces were affected by the racial or language condition they were in, we divided responses according to four stimuli groups: White Spanish faces (in all-White or all-Spanish conditions), White English faces (in all-White or all-English conditions), Black Spanish faces (in all-Black or all-Spanish conditions), and finally Black English faces (from all-Black or all-English conditions) (Fig. 3.5). A repeated measures ANOVA with response type (accurate or incorrect) as the factor with Experiment as the between-subject condition was conducted on White Spanish faces from the all-White condition (Experiment 1) and all-Spanish condition (Experiment 3). We found a significant main effect of response type ($F(1,61)= 52.612$, $p<0.001$) meaning there were more errors than accurate responses, and a marginally significant interaction of response type and experiment ($F(1,61)= 3.33$, $p= 0.07$). This was obtained by the high accuracy scores for the White Spanish faces in Experiment 3. This means that responses to White Spanish faces were affected by the language and racial surroundings. Simply, that responses to White Spanish faces differed if they were among White English faces or

Black Spanish faces. All other comparisons of similar stimuli across experiments did not reach significant results.

3.3 Discussion

The main objective of our study was to examine the interaction of language and race in creating social categories. We found that language categorization strength remains strong in own race (Experiment 1), other race (Experiment 2) and in mixed conditions (Experiment 5). Categorization by race was strongest in an L1 context (Experiment 3) but decreased in an L2 and mixed condition (Experiments 4 and 5), suggesting that race categorization is more malleable than language categorization. The following section discusses results based on language categorization, race categorization and in/out-group effects.

3.3.1 Language categorization

Our results support findings by previous studies that found that the categorization strength of accent was robust and remained strong in the single as well as multi-dimensional conditions (Pietraszewski & Schwartz, 2014a, 2014b; Rakić et al., 2011). While previous studies used accent differences (e.g. American vs. British English), we can conclude that different languages (Spanish and English) are also used in a similar way to create robust social categories. This result was expected, because like accent, language is a socially relevant and commonly encountered cue.

In addition, these previous studies examining the strength of accent categorization did so only in an own-race context (Pietraszewski & Schwartz, 2014a, 2014b; Rakić et al., 2011) while we also examined language categorization strength in an other-race context (Experiment 2). We found that while the strength decreases slightly compared to an own-race context, it remains a strong effect. This means that language categorization indeed is robust and reliable even in other-race contexts. This further supports the ELIT, in that language categorization remains robust regardless of the racial context.

3.3.2 Race categorization

On the other hand, race categorization was sensitive to the language environment. Race categorization was strongest when the two races were presented in a Spanish context (Experiment 3). When different race faces were presented in an English context (Experiment 4), categorization was the smallest among all experiments. This may indicate the malleability of racial categorization. However, it may also suggest that L2 language comprehension is a cognitively taxing stimuli, and therefore leaves less cognitive resources for social categorization. Therefore, in order to conclude that race is sensitive to linguistic context, it is also necessary to test racial categories within a native-language, but foreign-accented condition. Testing racial categorization in a foreign-accent L1 condition may remove the cognitive load of L2 comprehension, while allowing the examination of racial categorization strength in a non-native linguistic environment.

English is a widely spoken language by members of every race, and therefore it may be natural to expect and categorize White and Black faces within an English context. It would be interesting to examine whether other L2 contexts affect racial categorization in the same way. That is, if the L2 context in this study were Japanese, a monolithic group of people with more limited expectations of race, would White and Black racial categories be formed? (That is, participants would not expect White Japanese speakers and Black Japanese speakers and therefore may have difficulties creating social categories within an unnatural environment.) In this latter example, we would be examining not only the effect of L2 on overall racial categorization, but also the effect of specific racial expectations with regards to that language.

3.3.3 Interaction of race and language in social categorization

The main objective of this current study was to examine the interaction of language and race in forming social categories. Our results in totality do not provide a simple answer, but suggests the robustness of language categorization in the wake of racially diverse cues, while racial categorization seems to be a more malleable category boundary when confronted with other languages. This is in line with Pietraszewski's studies (Pietraszewski et al., 2014; Pietraszewski & Schwartz, 2014a) which found that the strength of language categorization remained stable even when coalitional information (charity group membership) was simultaneously available, creating Charity group A members with an American or British accent, and Charity group B members with an American or British accent (Pietraszewski & Schwartz, 2014a). Pietraszewski and colleagues (Pietraszewski & Schwartz, 2014b, 2014a) suggest that a mechanism dedicated to accent categorization evolved to keep track of speakers of different accents, which was a recurrent feature of ancestral environment. Our results further support this theory that language is a robust feature of social categorization that is not reduced when faced with other social cues, such as race. Here, our finding supports the ELIT (ethno-linguistic identity theory), in that linguistic information is a resilient and vital part of social memberships.

Previous studies (Kurzban et al., 2001; Pietraszewski et al., 2014; Pietraszewski & Schwartz, 2014a) found that race categorization was sensitive to coalitional cues. Our results support their findings, as racial categorization was also affected by cross-cutting language cues (Experiment 5). Previous research suggest that because human ancestors were much less likely to interact with others from another race, it is unlikely that humans evolved a system solely dedicated to track race. They instead suggest that categorization by race is in fact a byproduct of coalitional psychology, which evolved to keep track of coalitional information including patterns of cooperation or competition (see also Cosmides, Tooby, & Kurzban, 2003). That is, a mechanism dedicated to tracking and updating coalitional information is used to keep track of racial information, which in society may normally reflect different coalitional forces. However, Pietraszewski & Schwartz (2014a) also warn of the pitfalls of interpreting their findings as accent being a stronger or more important cue than race for categorization. Instead they argue that language and race correspond to different aspects of human 'groups' and can have different underlying mechanisms with complex relationships with one another.

While the memory confusion paradigm (MCP) may be a useful paradigm to explore social categorization, it may place heavy emphasis on linguistic cues. To explore whether this was true, a study examined whether the *fit* of the linguistic content to the face stimuli affected categorization strength (Karl Christoph Klauer et al., 2014). Fit was considered to be the degree of match between stereotypical expectations of the target's category (e.g. women talking about make up). They designed MCP experiments with two races (White and Black faces) crossed with different topics with varying fit: 1) basketball teams, giving rise to coalitions but unrelated to race, 2) racial equality, with high racial fit, and 3) neutral topics with no coalitional information or racial fit such as university life. The authors found that when these faces spoke on racial equality, categorization by race increased and was highest compared to when those same faces spoke about neutral topics such as university life or coalitional cues. This suggests the importance of the linguistic cue in the MCP and how it can manipulate racial categorization by precisely emphasizing or de-emphasizing race. It would be interesting to see whether emphasizing Spanish or English characteristics in linguistic content will increase language categorization as well. While our sentences were neutral sentences with no obvious racial or linguistic fit, studies such as the one mentioned above demonstrate the heavy reliance on the linguistic variable in a MCP experiment.

3.3.4 In/Out group effects

In addition to categorization strength, we examined recognition memory as an exploratory measure by comparing accurate responses to Spanish vs. English faces (in Experiments 1 and 2) and White vs. Black faces (Experiments 3 and 4). Results showed a higher accuracy for White faces compared to Black faces, signifying the advantage of a racial in-group member for memory, as has been consistently found in the past (Meissner & Brigham, 2001). No similar effect was found for language in/out-group accuracy responses suggesting the overall importance of race in driving in-group recognition advantages when compared to language membership.

However, the mixed condition paints a different picture. Black Spanish faces (double out-group) were recognized the most accurately. This is surprising, as according to previous studies on social factors that affect race recognition (Hehman et al., 2011; Shriver et al., 2008), double out-group members were never recognized more accurately than partial in-group members or double in-group members. First, it must be noted that the MCP paradigm is not classically used to examine in/out-group effects, and while results were significant, the margin of difference was below one response. This higher accuracy for Black English (double out-group) faces may be an expectancy effect. That is, Spanish participants may be more likely to expect Black faces to speak a foreign language than Spanish, and therefore may have been more likely to assign English sentences to Black than White faces, and Spanish sentences to White, rather than Black faces.

Every single face presented in this study had characteristics in 2 dimensions: race and language. Therefore all faces fell into one of four categories: 1) double in-group (White Spanish) 2) racial in-group/language out-group (White English), 3) racial out-group/language in-group (Black Spanish) or 4) double out-group members (Black English). The fact that accuracy for White Spanish faces (double in-group) was higher when among Black Spanish faces (Experiment 3) than White English faces (Experiment 1) implies the importance of the surrounding racial context on face recognition.

3.3.5 Summary and future directions

In sum, we examined the interaction of language and race in forming social categories, and while language categorization is robust and unaffected by racial contexts, race categorization may be more sensitive to language contexts. This supports previous studies examining language and race categorization in a MCP experiment.

In this chapter, we have established that language categorization remains stable, even in the face of other-race faces, while race is more sensitive to language cues. Therefore, the objective of the next chapter in this dissertation was to explore the effect of linguistic cues on other-race face perception and memory. Considering our current result suggesting that the demand of foreign language comprehension affected categorization, we used native- and foreign- accented Spanish sentences and Asian faces, which may reflect a more realistic environment the participants are generally accustomed to within the Barcelona context.

Chapter 4

The influence of native and foreign accent on other-race face perception.

4.1 Introduction

The face of a person holds many cues. They hold information as to the gender, age and race of a person, and can express a wide range of emotions. Using these cues, we categorize people into social categories. As we have discussed in the previous two chapters of this dissertation, we have shown that the language of a speaker influences the perception of their face, and that language and race interact in creating social categories. Our next question is whether linguistic cues (accent) influence the perception of other race faces. More specifically, can a native or foreign accent affect the perception of an other-race (Asian) face by assigning them as in/out-group members?

Decades of research have been dedicated to the perception and categorization of different race faces (Allport, 1954; Meissner & Brigham, 2001; Quinn, Lee, & Pascalis, 2018). The most popular method of examining race effects come from recognition studies in which own-race faces are better recognized than members of another race (for a review see Meissner & Brigham, 2001). While many of us may consider that a person's race is fixed and constant, many studies have looked at the top-down effects of categories on face-*race* perception (Freeman & Johnson, 2016; Freeman et al., 2011; Kim & Davis, 2010). Experimental studies have found that when classifying the face of a morph race, a face is more likely to be categorized as Black if they are wearing low socio-economic attire (Freeman et al., 2011). Even in observational studies collected from longitudinal databases, faces were more likely to be categorized as Black when they had been incarcerated (Penner, 2014), unemployed or impoverished (Penner & Saperstein, 2008), or had died by homicide (Noymer et al., 2011). These studies suggest that race can be affected by other social signals such as stereotypes or attitudes towards a certain racial group or certain characteristics.

This gives support for the dynamic-interactive model of face perception (see fig. 1.1 in the general introduction; Freeman & Johnson, 2016) in which face and race perception is not solely a bottom-up processing of visual cues, but that visual perception interacts with higher cognitive mechanisms such as stereotypes when viewing someone's face and race. Chapter 3 of this dissertation examined the ways language and race interact in creating social categories, and found that race categorization is affected by language. Hence, the question we would like to address in this study is whether the accent of a speaker (native or foreign) affects race *perception*. That is, does a foreign accent allow Asian faces to be perceived as more foreign, and does a native accent allow Asian faces to be perceived more like a Caucasian face?

According to the United Nations, the number of international migrants has continued to grow rapidly worldwide (United Nations, 2017). In particular for Barcelona, the largest immigrant population from non-European countries is from China (Ajuntament de Barcelona, 2017). Hence, it can be argued that Barcelona citizens are increasingly more accustomed to meeting Asians who speak Spanish. Larger immigrant population also suggests that there is an increase in children of immigrants born in Spain (who are not counted within immigrant population databases), who speak Spanish with a native accent. Modern, international societies allows for different races and ethnicities to live within a community with a varied set of accents and languages. Therefore it is becoming increasingly crucial to examine the ways linguistic cues and race interact when meeting such diverse peoples.

Compared to race perception studies, linguistic cues (language or accents) have only recently been used as a social cue in facial categorization (Baus et al., 2017; Pietraszewski & Schwartz, 2014a; Rakić et al., 2011). Therefore few, if any, studies have looked at the interaction of both race and language in creating social categories, and how they influence face perception. In the second chapter of this dissertation, we have observed that language categorization influences the visual perception of faces, and in the third chapter, we have observed that race categorization is influenced by language. Therefore, in this chapter, we aim to extend the findings of the previous chapters by examining the way *accents* influences the perception of other-race faces.

One study attempted to look at the effect of language and accent on race judgment and perception (Kim & Davis, 2010). In their study, computer generated faces were used, in which faces were morphed from Caucasian to Asian faces in 11-steps (i.e. they created a set of faces ranging from 100% Caucasian to 100% Asian, where the middle face was 50% Caucasian and 50% Asian). These computer-generated face-morphs accompanied either French or Japanese language (or French- and Japanese-accented English) and (Australian) participants were asked to classify them as Asian or Caucasian. In another condition, they were asked to classify their race in a silent condition. Results showed that faces were more likely to be judged as Asian when accompanying the Japanese cues when compared to the silent condition. However, the French cues did not affect racial classification. In simple terms, faces could be influenced by Japanese cues to become more Asian looking, but those same faces were not more likely to be classified as Caucasian even with the presence of French (cues). In a second experiment, Kim and Davis (2010) tested the same faces in a perceptual discrimination task, and found that only Japanese cues helped the discrimination between two faces, and only those which were racially ambiguous. French linguistic cues did not influence the discrimination of faces, ambiguous or otherwise. Therefore, they found that language affected the racial judgment and perception of racially ambiguous faces, and that only Japanese linguistic cues affected these faces in the direction of those faces becoming an out-group member.

In our experiments, we attempted to create a more natural setting, in which we used real face photographs and real audio recordings. In addition, we avoided having Caucasian voices accompany Asian accents. This may allow the viewing of believable faces when compared to morph faces and computer generated voices combined in unrealistic ways. Secondly, we used a paradigm in which low-level perceptual differences may be measured, and have been used to examine whether there are differences between races. We used the perceptual priming paradigm, similar to that conducted by Herrmann et al., (2007) in which two faces were shown one after the other very quickly, and participants distinguished whether they were the same or different. In their study, they found that reaction times differed between the two races, but only when the two faces were different. Simply put, when the two faces were the same, participants accurately and quickly judged that the faces were the same regardless of the race. When the two faces were different, participants were slower (but not less accurate) at determining that the faces were different, but only when they were Asian faces. They suggest that this reflects a deficit for Caucasian participants to process other-race faces. We therefore used this paradigm to see whether the perceptual difference could be affected by linguistic cues.

The aim of this study was to test whether the accent of a speaker affects face perception in a perceptual priming paradigm (Herrmann et al., 2007) and face recognition. We hypothesized that other-race faces would be categorized differently according to the accent they accompanied (native or foreign accent), and that this would lead to faster reaction times to Asian faces that accompanied a native-accent. By doing so, we examine whether other-race face perception is influenced by the social categorization by accent.

In order to do this, participants viewed three groups of photographs of female faces: Caucasians with native-accented Spanish, Asian faces with native-accented Spanish, and Asian faces with foreign-accented Spanish. Participants completed a priming task before and after seeing them with voices, to see whether the accent differentiated the two Asian face groups: the ones with a native-accent and those with a foreign-accent. Finally, we conducted a recognition test to examine whether memory for the two Asian face groups were affected by accent. We hypothesized that after the training session, reaction times to Asian faces with a native-accent will become more similar to the reaction times of the Caucasian faces, when compared to the Asian faces with a foreign-accent. As a control, for another group of participants, we conducted the same experiment but instead of voices, we used flags (Spanish or Chinese) as an alternative type of social cue.

4.2 Methods

4.2.1 Participants

We had a total of 69 participants (avg. 19.8 years old, SD=1.2 yrs, range 18- 25) who were recruited from the University of Pompeu Fabra database of students who signed up to participate in studies. Thirty-four participants (11 male) completed the experiment with speech in their training session, while 35 participants (15 male) completed the flag condition.

4.2.2 Stimuli

Asian faces were collected from an online database created for research purposes (CAS-PEAL-R1 face database, under the sponsor of the Chinese National Hi-Tech Program and ISVISION Tech. Co. Ltd, Gao et al., 2008). Caucasian faces were collected from several online, free face databases. Photographs of faces were in grey-scale, and faces were cut out in oval shapes so as to remove features such as face shape, hair styles, and other characteristics (see figure 4.1). In total, 100 Asian, and 50 Caucasian female faces were used, considering seen, new and filler faces.

Recordings were made using Audacity (v 2.0.3) at the University of Pompeu Fabra. Female native speakers of Spanish were instructed to read autobiographical sentences in a normal way, i.e.- normal speed, regular intonation. Foreign accented sentences were collected from Japanese students studying Spanish at the Kobe City University of Foreign Studies (Kobe, Japan). The average duration of foreign-accented sentences was 2588ms (SD=673ms), and the average duration of the native-accented sentences was 2393ms (SD=537ms), and this difference was significant ($t(178)=2.112$, $p=0.036$). Sentences had a range of 3 to 15 words, and the average number of words in the foreign-accented sentences was 6.5 words (SD=1.7 words), and the average number of words in the native-accented sentences was 7.7 words (SD=1.9), and this difference was also significant ($t(178)=-3.935$, $p<0.001$). However, as audio was only presented during the training phase, we believe this time difference did not affect perception or memory in the following procedures.

4.2.3 Procedure

Participants went through four phases in this study. First, the Priming Task, second the Training Phase, followed by the Post-training Priming Task and finally the Face Recognition phase. Faces were counterbalanced to create 6 lists. Counterbalancing was done in the following way: Asian faces were blocked into three groups, in which across three lists, they rotated whether they accompanied a native-accent, foreign-accent or appeared as filler faces, and Caucasian faces were blocked into two groups that appeared together as Caucasian-native or fillers. Finally, response keys were counterbalanced for all conditions to create a total of six lists.

In the first Priming Task, participants were shown two faces simultaneously, and asked to judge whether they were the same or different faces (see figure 4.1). A fixation cross appeared for 500ms, which was followed by the first face (S1) for 200ms, then a mask for 1000ms followed by the target face (S2), to which the participants answered was the same or different as the S1 by pressing on keys of the keyboard (keys *f* or *j*). There were 60 Asian faces and 40 Caucasian faces in this task. Twenty of those Asian faces and twenty of the Caucasian faces were used as filler faces for when the S1 and S2 presentation were different. The filler face was always in position S1. Forty Asian faces and 20 Caucasian faces underwent the following training session, in which they would be associated with an accent. All 40 Asian and 20 Caucasian faces were presented as both a 'same' condition and 'different' condition.

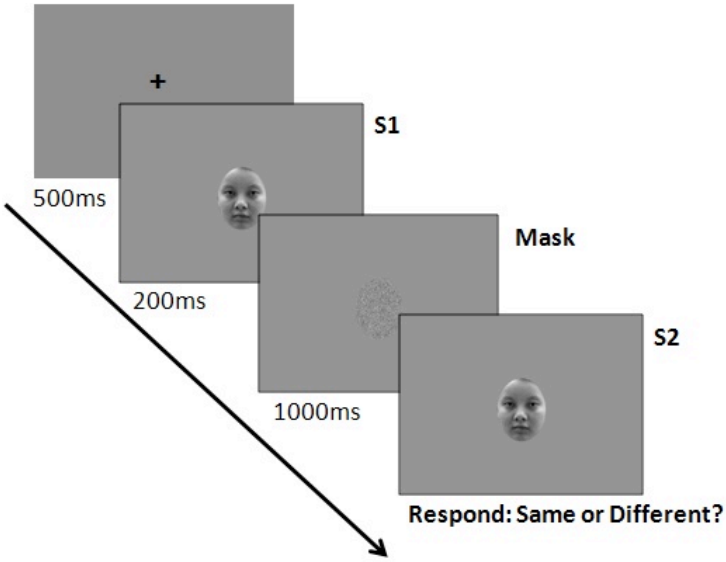


Figure 4.1 Procedure of Priming Task. Participants were presented with a 500ms fixation crossed followed by a quick presentation of the prime face (S1) for 200ms, which was followed by a mask for 1000ms. This mask overlapped the opening in which the faces appeared, which was filled with static. Then the S2 face appeared with the question of whether the S2 face was the same or different as the previously seen face (S1).

Once the Priming Task was finished, the participants continued into the Training Phase, in which they passively viewed faces accompanied with sentences. In this phase, a total of 40 Asian faces were presented, 20 of which accompanied native-accented Spanish accents and 20 of which accompanied foreign-accented Spanish. They were also presented with 20 Caucasian faces which all accompanied a native-accented voice. Each face was presented a total of three times during this phase, each time with a different sentence. Every face always accompanied the same voice, but with three different sentences.

After the Training Phase, the participants did the post-training Priming Task. This was identical to the first priming task, in that once again, they were instructed to respond whether the quickly presented faces were the same or different.

Finally, participants completed the Recognition Test, in which they were presented with all previously trained 60 faces (40 Asian, 20 Caucasian) mixed with another 60 new faces (40 Asian and 20 Caucasian faces), and asked whether they had previously seen the face or not (filler faces were not used in the recognition task). After a fixation cross for 500ms and a blank screen for 100ms, they were presented with the face and asked if they had previously seen this face by answering yes or no.

In a second experiment with a different set of participants, everything remained exactly the same except that the training phase accompanied flags instead of voices. Caucasian faces always accompanied a Spanish flag, while the Asian faces accompanied either a Spanish or a Chinese flag. The duration of presentation of each face was the same across both training sessions. This was done as a control experiment to see whether the results were affected by the number of presentations of the faces, and whether voices act as a unique cue.

4.3 Results

4.3.1 Priming Test

Audio (accent) group

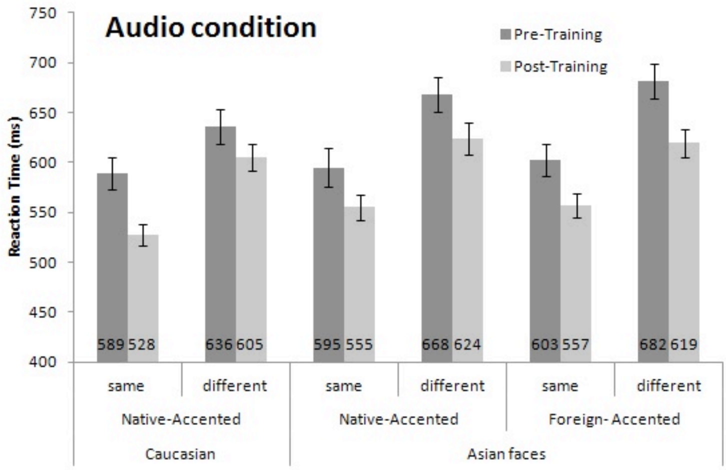
Reaction times and accuracy rates were collected from the S2 presentation of the face (see figure 4.2). Reaction times used in the analyses were only collected from accurate responses. While the faces in the pre-test had not been trained with an accent, we analyzed the faces in three groups (Caucasian with a native accent, Asian faces with a native accent and Asians with a foreign-accent), to collect the baseline reaction times for the three groups, especially between the two Asian groups (see figure 4.2). A repeated-measures ANOVA was conducted with three factors: training (2 levels: pre- or post- training), group (3 levels: Caucasian native, Asian native, Asian foreign) and presentation (2 levels: same or different S1/S2 presentations).

Results showed a main effect of training ($F(1,33)=21.39$, $p<0.001$) meaning that the pre-test was slower than the post-test. A main effect of group ($F(2,66)=16.08$, $p<0.001$) was found, and a post-hoc pairwise comparison, which used the Bonferroni correction for multiple comparisons, showed that the Caucasian faces were significantly faster compared to the Asian-native ($p=0.001$) and Asian-foreign accent ($p<0.000$). However, the two Asian face groups had the same reaction time ($p=1.0$). We also found a main effect of presentation ($F(1,33)=45.32$, $p<0.001$) revealing that participants were faster at same face presentations than different face presentations.

While we did not find the expected interaction between training and group ($F(2,66)=0.93$, $p=0.401$), an interaction between training, group and presentation was observed ($F(2,66)=4.513$, $p=0.015$). Posthoc pair-wise analyses revealed that there was an effect of race, in which reaction to Caucasian faces were faster than Asian faces for the different presentations in the pre-training test ($p<0.046$), and in the same presentation in the post-training test ($p<0.006$). Unfortunately, reaction times to the two Asian face groups were exactly the same ($p=1.0$), regardless of training session.

The same repeated-measures ANOVA was conducted for accuracy rates (three factors: training, group and presentation). Results showed a main effect of training ($F(1,33)=24.821$, $p<0.001$) in which the accuracy rate improved in the post-training test (see figure 4.3). We also obtained a main effect of group ($F(2,66)=6.95$, $p=0.002$). These results were revealed to be an effect of higher accuracy scores for Caucasian faces than both Asian faces ($p<0.009$) in a post-hoc analysis. The two Asian face groups had the same accuracy score ($p=1.0$), and there was no interaction between group and training ($F(2,66)=0.09$, $p=0.914$) meaning that accent training had no effect on the accuracy scores of the two Asian face groups.

In sum, the priming experiments in the voiced condition revealed a main effect of race in which participants reacted faster and more accurately to Caucasian faces than to Asian faces. However, there was no effect of accent training on the Asian faces on perception as was expected.



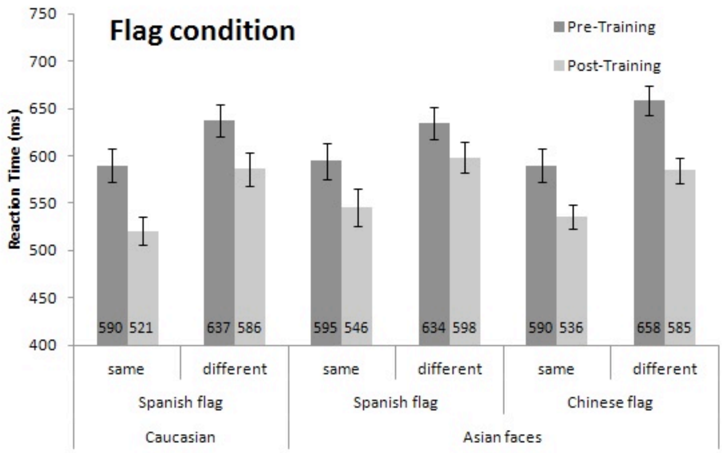


Figure 4.2 Reaction times for priming experiment in both pre- and post-training sessions. Top graph shows results from the audio (accent) group, while the bottom graph shows the reaction time for the flag group. All error bars reflect standard error.

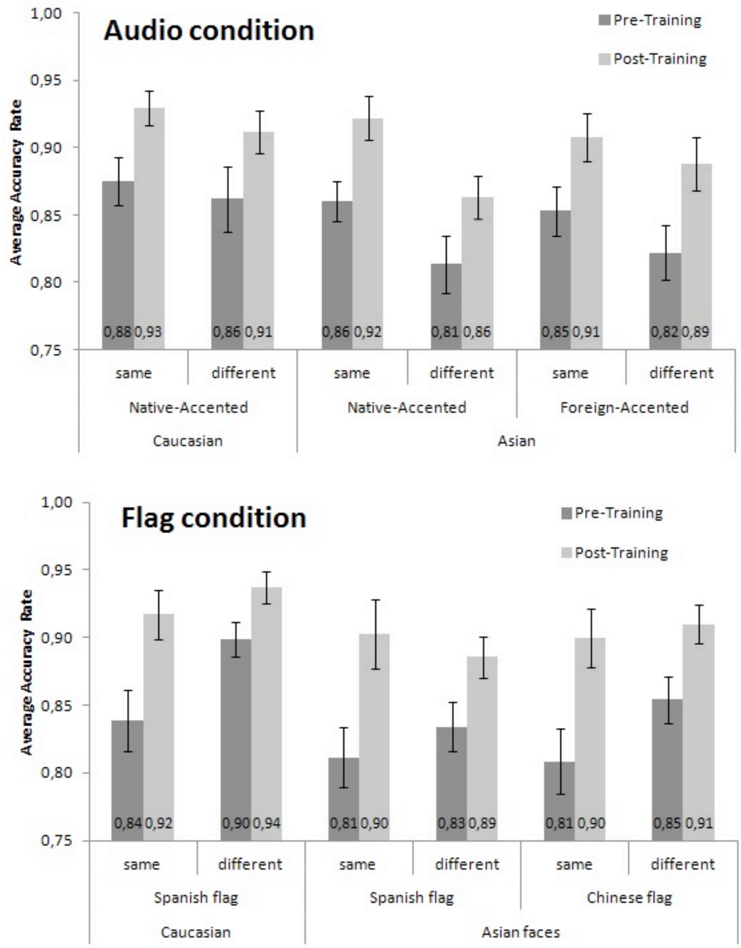


Figure 4.3 Average accuracy rates for Pre- and Post-training Priming tests in the Audio (top) and Flag (bottom) groups.

Flag Group

The same repeated-measures ANOVA was conducted for the prime-target task in the flag condition for reaction times. The analysis had three factors: training (2 levels: pre and post training tests), group (3 levels: Caucasian with a Spanish flag, Asians with a Spanish flag and Asians with a Chinese flag) and presentation (2 levels: same or different presentation).

We found a main effect of training ($F(1,34)=33.54$, $p<0.001$) in which the participants were faster in the post-training session (see figure 4.2), and a main effect of presentation ($F(1,34)=50.732$, $p<0.001$) in which the responses were faster to the same face presentation (563ms) than the different face presentation (617ms). We found no main effect of group ($F(2,68)=2.114$, $p=0.13$), meaning that there were neither differences between the Caucasians and Asians, or a difference between the two Asian groups that had been trained with Spanish or Chinese flags. This means that we did not observe an other-race effect in the reaction times for the flag-trained group, and that flag-training had no influence on the Asian faces. The interaction between training and group approached the level of significance ($F(2,68)=2.52$, $p=0.088$), but this was driven by the effect of training on all groups ($p<0.003$). No other interactions reached significance ($p>0.109$).

The same analysis was done for accuracy rates (see figure 4.3). Results showed a main effect of training ($F(1,34)=24.40$, $p<0.01$), in which accuracy rates improved in the post-training test. We also found a main effect of group ($F(2,68)=13.37$, $p<0.001$), and a post-hoc pair-wise comparison showed that accuracy rates were significantly higher for Caucasian faces when compared to Asian faces which were trained with Spanish flags ($p<0.001$) and Chinese flags ($p=0.003$). The accuracy rates for the two Asian groups were the same ($p=1.0$). We also did not find an interaction between training and group ($F(2,68)=0.562$, $p=0.554$) meaning that while we found a main effect between the two races, there were no differences between the two Asian faces due to flag exposure. The only interaction that approached a level of significance was the interaction between group and presentation ($F(2,68)=3.096$, $p=0.052$), and a post-hoc analysis showed that this was driven by the race effect (Caucasian versus both Asian groups, $p<0.017$) when faces were different. All other interactions were not significant ($p>0.083$).

In summary, results from the priming experiment in the flag-trained group only revealed a main race effect between Caucasians and Asian faces in accuracy, but not reaction time, and there was no effect of flag training on the two Asian faces.

4.3.2 Recognition Task

Audio (Accent) Group

After the post-training prime-target task, participants were shown all previously seen faces as well as new faces, and asked whether they had seen the faces before. Accuracy scores were calculated instead of d-prime scores, because d-prime scores could not be calculated for two Asian face types (native or foreign accented) with only one set of 'new' faces. That is, there were 40 new faces and 40 seen Asian faces, but 20 of those were native-accented and 20 were foreign-accented, and therefore d-primes of the two Asian groups could not be calculated.

Accuracy scores were analyzed in a repeated-measures ANOVA for the five different groups of faces: new Caucasian faces, old Caucasian faces with native accents, new Asian faces, old Asian faces with native accent, and old Asian faces with foreign accents (see figure 4.4). There was a significant effect of group ($F(4,132)=22.99$, $p<0.001$), and a post-hoc pairwise analysis showed that accuracy scores were higher for the new Caucasian and new Asian faces when compared to all old faces ($p<0.02$). Additionally, a race effect between the two new faces was found, in which the new Caucasian faces were more accurately responded to (0.88, SD=0.15) than new Asian faces (0.75, SD=0.18), and this effect was significant ($p=0.001$).

Most importantly, accuracy scores for Asian-native (0.57, SD=0.17) and Asian-foreign accented faces (0.56, SD=0.17) were the same ($p=1.0$). Asian-native approached significance when compared to chance ($t(33)=2.03$, $p=0.051$), and Asian-foreign was significantly different from chance (t(33)= 2.06, p=0.047), meaning that their performance was above chance level. Therefore, race effects were found only for new faces, and accents did not affect the recognition accuracy between Asian seen faces.

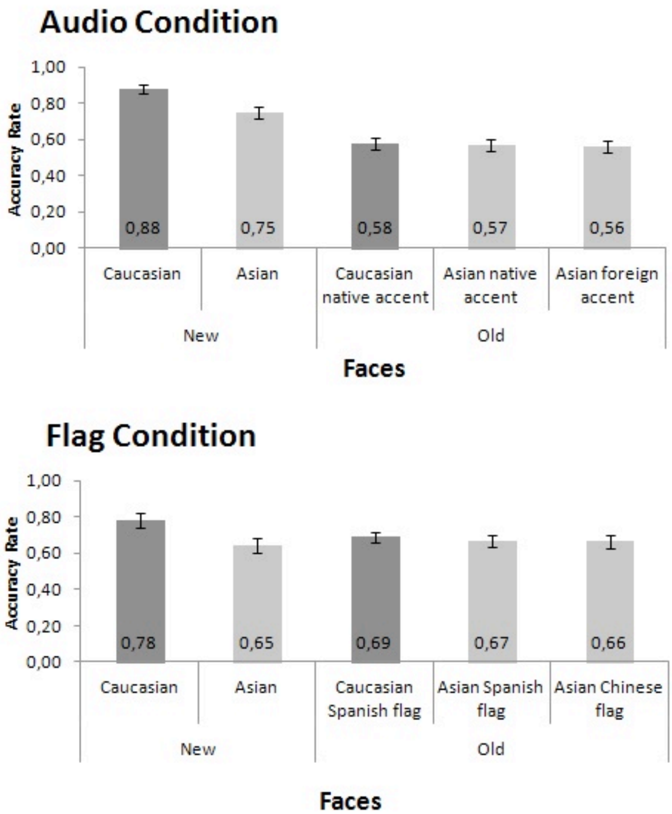


Figure 4.4 Recognition Accuracy for both audio (top) and flag (bottom) condition. Error bars denote standard error.

Flag Group

Accuracy rates were also collected for the recognition task for faces that were trained with Spanish or Chinese flags (see figure 4.4). A repeated-measures ANOVA was conducted to see whether accuracy rates were different among the five face groups (new Caucasian, old Caucasian-Spanish flag, new Asian, old Asian-Spanish flag, and old

Asian-Chinese flag). Results showed no significant effect of face group ($F(4,136)=2.0$, $p=0.149$) meaning there were no differences between race, flag training or between old and new faces. Accuracy rates for all faces were above chance level ($p<0.001$), meaning the participants had been performing their task.

4.3.3 Summary of Results

In sum, we found a general other-race effect (ORE) in both the pre- and post-training for both the voiced and flag conditions. For the group that received voiced-training, this ORE was reflected as faster reaction times and higher accuracy scores for the Caucasian faces than the Asian faces, while for the group that received flag-training, this difference was only reflected in accuracy scores.

The ORE was also reflected in the higher recognition accuracy for the new Caucasian faces when compared to new Asian faces, but only in the voice-training group.

Most importantly, the perception and recognition of Asian faces was not affect by accent or flag training.

4.4 Discussion

By using a priming and face recognition paradigm, we attempted to influence the perception and recognition of other-race Asian faces by training them with either native- or foreign-accented Spanish voices. We found faster and more accurate responses to Caucasian faces than to Asian faces, but we did not obtain differences between the two Asian faces that had been trained with different accents.

In the priming experiment, we consistently found differences between the Caucasian and Asian face perception in a general other race effect. These results replicate previous studies which also found the other-race effect in perception (Herrmann et al., 2007). The results from our pre-training test is most comparable to their results, as they did not include a post-training test. While Herrmann et al., (2007) only found differences in reaction times, our results showed a broader race effect in which race differences were found across pre- and post-training sessions in both reaction time and in accuracy.

Race differences persisted in the post-training test, meaning that the other-race effect was not reduced by the presentation in the training. While the training session was not an explicit training period with instructions to memorize the faces, this points to the robustness of the other-race effect on perception, and its persistence even after exposure to those faces in a training session, regardless of training type.

We also found race effects in the recognition tests where accuracy scores were higher for *new* Caucasian faces compared to *new* Asian faces. Here we replicate decades of research finding poorer recognition for other-race faces when compared to own-race faces (e.g. Allport, 1954; Lindsay, Jack, & Christian, 1991; Meissner & Brigham, 2001). In a meta-analysis of over 39 studies on the other-race effect, Meissner & Brigham, (2001) found that participants were more than twice as likely to accurately identify own-race faces as new vs. old, when compared to the other-race faces. Our data replicate these findings in that new Caucasian faces were more accurately identified as new compared to new Asian faces.

However, we did not find race differences between the *old* (seen) faces, which may reflect an effect of general training. In the same meta-analysis study, Meissner & Brigham, (2001) found that old own-race faces were more correctly identified as old than other-race faces. Our study did not find this effect, and therefore may indicate an effect of training on seen faces. Previous studies have provided evidence that explicit training can reduce the other-race effect in face recognition (Elliott, Wills, & Goldstein, 1973; Heron-Delaney et al., 2011; Lebrecht, Pierce, Tarr, & Tanaka, 2009; Malpass, Lavigueur, & Weldon, 1973; Meissner & Brigham, 2001; Tanaka & Pierce, 2009). For example, Tanaka & Pierce, (2009) found that individuation training affected the rate of hits (seen faces correctly identified as seen), but not the correct identification of new faces as new. Therefore, our results may reflect an effect of training on the recognition of seen faces.

Interestingly, recognition accuracy was higher for the flag-training group than the voice-training group. As the flag-training group was used as a control experiment, this reflects a) that participants did the experiments accurately, and b) flag-training experiment may have been easier than the voice training group. While the flag-training only had two cues, the voice condition had two accent types, which included different voice identities and different sentence contents. This may have increased the cognitive demand for the participants during the training period, making subsequent recognition more difficult.

The main goal of this study, which was to examine the effect of accents on the perception of other race faces did not yield significant results. That is, we were unable to obtain differences between the Asian faces that accompanied native or foreign-accented sentences in perception or recognition. This may be due to the inability of accents to influence the perception and recognition of other-race faces.

Previous studies on the effect of linguistic cues on face race categorization and perception revealed that racially ambiguous, morphed-faces were more likely to be categorized and perceived as Asian when accompanying Japanese language or accent (Kim & Davis, 2010). However, the same racially ambiguous faces were not more likely to be categorized as Caucasian when they accompanied French (language or accent). Furthermore, neither language affected the categorization of unambiguously (100% morph) Asian or Caucasian faces. Therefore, the effect of language on other-race face perception may only act in two ways. First, that language only affects the perception of racially ambiguous faces, and second, that this only occurs in the direction of those faces becoming a racial out-group from the participants.

In the second chapter of this dissertation, we established that language affects early stages of face perception of own-race faces. In our third chapter, we revealed that language is a robust social category while racial *categorization* can be modified by language. However, we could not yield significant results indicating that language (accents) affects the visual *perception* and recognition of other-race faces. Therefore, language effects on face perception may be limited to pre-attentive visual perception that does not alter downstream behavioral effects (such as reaction time, accuracy or recognition). Future studies may consider the use of ERP and other neuro-cognitive techniques in examining whether language affects other-race perception.

Our null findings may also be due to a limitation in our design. Compared to the previous two studies in this dissertation which used eight faces, this study compared 40 faces trained with voices (or flags). This may have been a very cognitively demanding quantity of faces, especially when also viewing another set of Caucasian faces. Perhaps having different races was a much more salient cue than accent differences within a race, and therefore training could not overcome the general other-race effect. In addition, studies which looked at the effect of training or instructions to reduce the other-race effect administered a much more explicit and intense training session (Bornstein et al., 2013; Elliott et al., 1973; Lebrecht et al., 2009; Malpass et al., 1973; Meissner & Brigham, 2001). Therefore, future studies examining the role of accent and languages on the perception of other-race faces should consider longer or explicit training sessions.

In conclusion, while we obtained general race effects between Caucasian and Asians, we did not find an influence of accent on the perception of other-race faces. As immigration and internationalization allows for more diverse peoples with different racial, ethnic and linguistic backgrounds, it is increasingly crucial to investigate the ways linguistic cues and faces interact in the ways we see, and also ultimately behave towards them. More studies should be conducted on these socially relevant parameters and how they influence each other.

Chapter 5

General Discussion

Past studies have examined the ways we categorize people, and importantly for our purposes, how people categorize others on the dimension of race and language. Race as a social category has been extensively studied, and its effect on person recognition may even be familiar to the layperson who has experienced the other-race effect. However, effects of *language* on social categorization is perhaps less obvious. Recent research has shown that language is indeed a category of social categorization (Pietraszewski & Schwartz, 2014b, 2014a), and may even have consequences such as the other-accent effect (Stevenage et al., 2012), in which participants were better at recognizing voices from their own-accent than an other-accent. Therefore, race and language are both used in social categorization, and this categorization may affect the way we remember and interact with others.

What has become more prevalent in recent years, is the examination of ways different social cues act *together* in influencing categories, judgments, impressions, memory, etc. (Freeman & Johnson, 2016; Hansen et al., 2017; Johnson, Freeman, & Pauker, 2012; Pietraszewski & Schwartz, 2014a; Rakić et al., 2011). What has not been explored in such studies is the question of how language and race interact in creating social categories, and how it affects the way we *perceive* them. Therefore, this dissertation aimed to tackle such questions as to what happens when different race and different language cues coexist in the people we meet.

5.1 Summary of findings

The main aim of this dissertation was to examine the influence of language on social categorization, and its effect on face perception and recognition, including other race faces. We first examined the effect of language on social categorization of faces, and whether this categorization affected early visual perception of those faces in an own-race context (Chapter 2). We then extended this finding by investigating whether language interacts with race in creating social categories. Chapter 3 explored whether language is used as a social category in own- and other-race faces, and whether race is used as a social category in a native- and foreign-language context. We then tested how language and race interact in creating social categories in a multi-group condition when both cues were available. Finally in Chapter 4, we sought to investigate whether native- and foreign-accents influence other-race face perception and recognition.

In **Chapter 2**, we replicated findings by other researchers examining the role of linguistic cues on social categorization as well as the effect of categories on early visual perception. Behaviorally, our results show that language is used as a social category for faces, replicating previous findings that *accent* is used as a dimension of social categorization (Pietraszewski & Schwartz, 2014b; Rakić et al., 2011). We further examined whether the social categorization of faces by language affects early visual perception of those faces, and indeed found that similar to previous studies on color perception, they do (Clifford, Holmes, Davies, & Franklin, 2010; Thierry et al., 2009). We add to the growing body of literature finding that categories (e.g. color, gender) affect the early stages of visual perception, even in complex visual stimuli such as faces.

The contributions of this chapter in the literature supporting top-down effects of category on visual perception are considerable.

Firstly, previous studies on the effect of categories of faces on visual perception have not been tested purely in a top-down manner. That is, categories such as gender or emotions are established in a bottom-up manner (Czigler, 2014). In contrast, language categories in our study cannot be detected visually from the faces. Therefore, we provide evidence that non-visual categories affect early visual perception in a top-down manner.

Secondly, the categories were newly, and implicitly learned in the laboratory in a short amount of time. This allows future studies to test for categorical effects on visual perception learned in the lab. To our knowledge, two studies have attempted to train participants with newly created categories and tested its effects on visual perception in an oddball paradigm (Clifford et al., 2012; Yu et al., 2017). Categories of color only affected post-perceptual stages of processing (Clifford et al., 2012) while face categories affected early and late perception (Yu et al., 2017). Our results match those of Yu et al., (2017) in that categories were detected at the early stages of face processing, but the patterns of category detection in later time windows were different. Our study provides more evidence for the effect of categories on early stages of visual perception, however, more studies are required to understand the effect of categories on later time windows. One possibility for the discrepancy between our studies might stem from the fact that in contrast to most studies using the oddball paradigm, our paradigm used a convoluted design in which there were more than one standard stimulus and deviants from both within and between the categories. Nonetheless, we obtained categorical effects in early visual perception. This method allows for the distinction between the detection of lower-level differences in the visual stimuli (e.g. face A vs. face B) and the detection of different categories from the faces (i.e. group A's category vs category B's category). Therefore, we expand the literature on newly learned categories on visual perception and suggest that more investigation is necessary.

In **Chapter 3**, we examined the interaction of language and race in creating social categories, especially when they are both presented to create multi-group members. As mentioned repeatedly, several studies have shown that accents are a dimension of social categorization (Pietraszewski & Schwartz, 2014b, 2014a; Rakić et al., 2011), and our study provides further evidence of the importance of *language* in such processes. What is novel in our study is the examination of language categorization within our own- and other-race situations, as well as race categorization in our native- and foreign- languages (Experiments 1, 2, 3 and 4) as well as the interaction of language and race in creating social categories (Experiment 5). We found that language categorization remains relatively stable across own- and other-race conditions, and in crossed-dimensional conditions. However, we found that race categorization was more malleable across language conditions and when it was crossed with language. Therefore, we conclude that language is a strong category of social categorization even in the face of other races, while race is perhaps a more delicate category cue affected by language contexts.

To our knowledge, our study is the first to examine the effect of language and race when they are crossed to create multiple-group members. Previous studies have crossed race with other cues, such as coalitional groups, sex, political affiliation (for a list see supplementary material in Bor, 2018) and have found that race is a malleable cue when crossed with other social information. Our study adds to this body of literature in suggesting that language also affects race categorization.

As the second and third chapters established that language affects face perception, and that language and race interact in creating social categories, the natural progression was to examine whether linguistic cues affect visual perception of other-race faces. This was our objective in **Chapter 4**. We examined the effect of native- and foreign-accented speech on face perception and recognition.

Previous studies found that the Japanese language or accent made racially ambiguous face-morphs more likely to be categorized and perceived as Asian. In this study, we used a perceptual priming task and face recognition test to examine whether native- and foreign-accented Spanish sentences influenced the way Asian faces were perceived and later recognized. We replicated a general other-race effect in which participants performed better for Caucasian faces than for Asian faces in perception and recognition tests (Herrmann et al., 2007; Meissner & Brigham, 2001).

However, we obtained no differences among the Asian faces that had been trained with native- and foreign-accents. We argue that while this may be due to a limitation in the design of the experiment, it is also possible that accent cues cannot affect the perception of other-race faces, at least in behavioral measures. We propose the necessity of future studies to examine the effect of accents and language in people perception.

In summary, through this dissertation, we have established that language is used as a cue for social categorization, and that this language categorization affects early stages of face perception. Secondly, we provide evidence that language is a robust cue for social categorization, even in different racial contexts, while race is a more sensitive and malleable cue that can be influenced by language contexts. Finally, we did not obtain an influence of native and foreign accents on other-race face perception. However, when considering the findings of the first two studies, we strongly recommend the further examination of the effect of language and accent on the perception of other-race faces.

5.2 Methodological Considerations of Dissertation

5.2.1 WEIRD Bilingual participants

All participants in this dissertation were WEIRD participants. That is, they were Caucasian university students from a Western, Educated, Industrialized, Rich and Democratic (WEIRD) society (Henrich, Heine, & Norenzayan, 2010a, 2010b). While most psychological studies, including those on race, have been done on WEIRD participants, authors tend to generalize finding to the general human population. However, recent studies have found that in fact, these particular populations are not the standard of human behavior, but perhaps even outliers with respect to the world population (Henrich et al., 2010b).

For example, one study compared the eye movements of Western Caucasian and East Asian (Chinese) participants when viewing a face. Researchers found that while Caucasian participants showed a triangle-pattern of fixation for faces focusing on the eyes, East Asian participants fixated on the central region of the face (Blais, Jack, Scheepers, Fiset, & Caldara, 2008). Authors therefore provide evidence that social experience and culture affects the way people view faces. Hence, it is possible that our findings from Chapter 2 may be different across cultural settings. That is, language categorization's effect on face perception may be altered due to cultural differences in general face processing.

In addition to the general way we process faces, the racial and cultural context differs from society to society, and therefore we may perceive race and social groups more generally, differently across cultures. Several studies have found that participants from different ethnic, cultural or national backgrounds show different patterns of recognition to out-group members when compared to Western participants (Chiroro, Tredoux, Radaelli, & Meissner, 2008; Ng, Steele, & Sasaki, 2016; Sporer, 2001). For example, Ng and colleagues (2016) compared the recognition accuracy for Caucasian and Asian faces by Caucasian-Canadian and East Asian participants (living in Canada). They also added an additional dimension of in/out-group team membership. While Caucasian participants showed better recognition towards own-race faces and in-group members, East Asian participants did not. Authors suggest that because face recognition is, in part, based on social categorization, it depends on how different cultures create social categories. That is, while Western (especially North American) cultures consider social groups to be represented as broad social collectives, East Asian cultures consider social groups to be a network of those with direct or indirect *personal* ties (Brewer & Yuki, 2007; Yuki, 2003, 2011). Therefore, such cultural differences need to be addressed in future studies.

Additionally, experiences with own- and other-race faces will inevitably be different for people from the racial majority compared to those from the racial minority. While a person from the racial majority will often meet others from their own-race group (because they are the majority race), they may have minimal contact with those from the other-race group. On the other hand, someone from the racial minority group will most likely have contact with the majority other-race group as well as own-race minority members. Therefore, the other-race effect cannot be assumed to be a simple reversed effect when testing participants from different races.

In addition to the cultural background of the participants, most of the participants in our studies were early Spanish-Catalan bilinguals who also spoke English as a foreign language (some participants spoke other additional languages). This may have impacted our studies in two ways.

First, there have been some studies showing that multi-linguals are more open-minded to out-group individuals (Mepham & Martinovic, 2018) or that bilingualism affects the ORE (Burns, Tree, Chan, & Xu, 2018; Kandel et al., 2016). In Kandel et al., (2016), researchers found that early bilingual participants (same population as in current study) did not exhibit an ORE. In addition, Burns and colleages (2018), found that while their bilingual population did exhibit an ORE, the magnitude of the effect decreased as proficiency in the second language increased. That is, participants with a higher proficiency in their second language exhibited the ORE to a lesser degree.

The experiments in Chapter 4 expand on these studies, as our participants did show the ORE, replicating the general ORE found in Burns et al., (2018). However, if bilingualism shapes the *magnitude* of the ORE, this may also explain the results found in Chapter 3. The Memory Confusion Paradigm (Who Said What? paradigm) does not measure the ORE in a direct way. However, if proficiency of a second language modulates the recognition of other-race faces, it could also partially explain the flexibility found in the categorization strength of race across the experiments. While studies on the effect of bilingualism on the ORE are unclear, it is worth noting that studies have found mixed results, and therefore future studies must consider the linguistic background of their participants when conducting studies with a racial dimension.

Secondly, it is feasible that bilinguals (especially ones who live in a heavily bilingual context) are more likely to assume that other speakers are also bilingual. This may lead them to consider someone who speaks a foreign language as a bilingual. This consideration may have consequences on how bilinguals categorize others by language. Simply put, monolinguals may be more likely to use language as a strong social category (e.g. they don't speak my language), whereas for bilinguals, this cue may be more flexible (they don't speak my language *now*). Results from Chapter 4 show that our bilingual participants did indeed categorize others by language (Spanish and English), and this was a robust effect. Future studies may consider the ways bilinguals categorize speakers from their two native languages, for example with Spanish and Catalan, which may act in different ways compared to a native versus foreign language.

Furthermore, these combined results point to the necessity to differentiate results obtained from paradigms measuring social categorization, perception and recognition. This will be discussed in the following section.

Due to the aforementioned effects of cultural, racial and multilingual contexts, we must be careful when generalizing results as a broad representative of all peoples. Different studies must be conducted across different cultural, racial and linguistic environments using different races, languages, or accents in order to have a better understanding of the interplay of such factors in social interactions.

5.2.2 Methodological considerations

As mentioned in the previous section, we must differentiate our interpretation of evidence obtained from experiments designed to measure social categorization, perception and recognition. That is, if faces are categorized along a certain dimension (for example in a Memory Confusion Paradigm), it does not necessarily follow, that this category will always affect face *perception* and *recognition* over other cues.

In specific terms, Chapter 3 showed that language categorization was a robust dimension of social categorization across different racial contexts, and that racial categorization was altered in different linguistic contexts. If this were to be interpreted as language being a better cue for face memory than is race, this could mean that other-race members who speak in a native language (relative to the viewer) could be better recognized than own-race members who speak with a foreign-language. Failure to obtain significant differences between the Asian faces that were trained with accents in Chapter 4 limit us from making conclusive statements on this matter. Our null findings may be due to limitations of the design, but it is also possible that face recognition and the other-race effect is not driven solely by social categorization processes, supporting hybrid theories of perceptual expertise and socio-cognitive models in face recognition (Young et al., 2012).

We have also used two methods to measure face *perception*. In Chapter 2, we used event-related potentials (ERPs) to measure early visual perception, whereas in Chapter 4, we used the priming paradigm to examine differences in reaction time. Results from Chapter 2 showed that language categorization affects early stages of face perception, but we were unable to find perceptual differences (i.e. reaction times) to faces trained with different accents in Chapter 3. Therefore, it is possible that language's effect on face perception is measured more easily with ERPs than by behavior. Conducting further studies to answer these questions may be necessary.

5.3 Social Relevance, Implications and Future Directions

5.3.1 Social Relevance of Dissertation

As mentioned several times throughout the thesis, it is increasingly becoming more likely that we meet diverse groups of people from different races who speak many languages and come in all combinations. Stereotypes of what a "native speaker" looks like will constantly be challenged. Race will become a less reliable cue for someone's nationality or language background, while language will also become a less reliable indicator of how that speaker may look like.

With immigration and globalization on the rise, we cannot ignore the current reality of a multi-ethnic and multi-linguistic environment. While we will definitely meet someone from another race or experience someone speaking to us in our foreign languages, it is just as likely that *we* become the 'outsider' when we travel, live abroad or find ourselves in multicultural settings. Therefore, this is an issue that affects each and every one of us with a face and a language. Results from such diverse studies may have far-reaching consequences such as the use in eye-witness identification, labor market discrimination, and language and racial integration policies of schools and governments.

More studies should be conducted in different cultural, racial, linguistic and social environments considering the diverse backgrounds and experiences of the participants in each context. As mentioned previously, the societal context in which these experiments are conducted may affect results in a myriad of ways. Nevertheless, this line of research is a much needed and relevant study, and therefore researchers should not be discouraged from pursuing similar or replication studies across different cultural contexts.

5.3.2 Color Blind Racial Ideology in Experimental Psychology

Color-blind racial ideology (CBRI) is a popular philosophy which stresses deemphasizing, minimizing, or ignoring intergroup, racial distinctions for advancing racial equality (Apfelbaum, Norton, & Sommers, 2012; Neville, Awad, Brooks, Flores, & Bluemel, 2013). The appeal of such approach to increasing diversity is understandable. If discrimination occurs due to the perception of racial differences, then removing or becoming "blind" to such perceptual differences should also remove or limit downstream consequences of racial prejudice. Simply put, if one cannot perceive any differences, one can not discriminate along those differences.

When reviewing the literature on social categorization by race, it is often met with the ideology that if people stop categorizing others by race or stop encoding and recalling the race of others, it will ultimately lead to less racial tension and promote a color blind society (Kurzban et al., 2001; Rakić et al., 2011).

> *e.g.*
> *Given that categorizing people into groups along nearly any dimension elicits discrimination, it would*

> *be discouraging to learn that the human mind was designed such that people cannot help categorizing others by their race. This would imply that racism is intractable.*
> *...the prospects for reducing or even eliminating the widespread tendency to categorize persons by race may be very good indeed.*

-Kurzban, Tooby, & Cosmides, 2001

Despite the prevalence of the CBRI, recent evidence from psychological studies have questioned the advantages in having such an approach to diversity and race relations. In fact, many studies have shown that not only is the approach ineffective, but that it promotes interracial tension and potential inequality (for a review see: Apfelbaum et al., 2012; Neville et al., 2013). One such study tested White participants in a Political Correctness Game in which participants were required to describe other individuals. Participants were less likely to use race as a descriptor when paired with a Black partner, compared to a White partner. Not only did this impair communication, but it also led to the White participants making less eye contact with the Black partner, ironically resulting in being perceived as less friendly (Norton, Sommers, Apfelbaum, Pura, & Ariely, 2006). Therefore, not mentioning or ignoring racial characteristics can inhibit smooth inter-racial interactions.

Recent work has also suggested that deemphasizing race in an effort to integrate racial diversity can shape an individuals' attitude towards racial out-groups. Studies have found that participants exposed to color-blind racial ideology later displayed greater degrees of both explicit and implicit racial bias (Holoien & Shelton, 2012; Jackson, Wilde, & Goff, 2016; Richeson & Nussbaum, 2004; Wolsko, Park, Judd, & Wittenbrink, 2000). These studies together suggest that CBRI not only prevents friendly interracial interactions, but that ironically, even with the best of intentions, it can increase or facilitate racial prejudice.

As studies examining the effect of top-down knowledge on the categorization of race suggests (Freeman & Johnson, 2016; Freeman et al., 2011; Kim & Davis, 2010; Noymer et al., 2011; Penner, 2014; Penner & Saperstein, 2008), the perception of race is not limited to a bottom-up perception of racial differences. As described in the dynamic-interactive model (see figure 1.1), facial features (e.g. thinner eyes, darker skin) feed social categories, which activate stereotypes and attitudes, but those attitudes and stereotypes in turn affect the categorization and perception of those faces in a top-down manner.

Therefore, the goal of studies in examining race as a social category should not be to erase, over-ride, or make race irrelevant with the intervention and help of other social cues. Through this dissertation, we emphasize that language did not *remove* the categorization of race (Chapter 3) or that language could be used to eliminate the perception of race (Chapter 4). What is proposed to be an effective alternative to the CBRI, is *multiculturalism,* an approach in which racial differences is acknowledged, recognized and celebrated. Empirical evidence support that a multicultural approach is more effective in promoting racial diversity, and is favored by racial minorities (Apfelbaum, Sommers, & Norton, 2008; Plaut, Thomas, & Goren, 2009; Richeson & Nussbaum, 2004). Therefore, future studies should promote a multicultural approach when interpreting and upholding their findings when using race and other social categories.

5.4 Concluding remarks

What do we *see* when we meet diverse people? Do we see their language? and what happens when their language conflicts with their race? This, in essence, was the main question of this dissertation. We explored this question by examining the ways language and race interact in creating social categories, and how this categorization affects how we perceive them. We found that language is used in social categorization, and this categorization affects early and late stages of face processing. When language and race are both available cues in social categorization, language remained as a robust cue, while racial categorization was altered by linguistic contexts. Finally, when examining whether native and foreign accents affected the perception and recognition of other-race faces, we were unable to find a significant effect of accent, but a general effect of race.

Findings from this dissertation enrich and expand our knowledge of how language and race interact in the perception of other people with fruitful data and socially relevant considerations. Internationalization, immigration and increased travel makes this a socially relevant topic now more than ever, and future multicultural approaches should be considered.

References

Ajuntament de Barcelona. (2017). La població estrangera a Barcelona. *Departament d'Estadística. Gerència Adjunta de Projectes Estratègics*, 1–139.

Allport, G. W. (1954). The Nature of Prejudice.

Apfelbaum, E. P., Norton, M. I., & Sommers, S. R. (2012). Racial Color Blindness: Emergence, Practice, and Implications. *Current Directions in Psychological Science*. https://doi.org/10.1177/0963721411434980

Apfelbaum, E. P., Sommers, S. R., & Norton, M. I. (2008). Seeing Race and Seeming Racist? Evaluating Strategic Colorblindness in Social Interaction. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/a0011990

Athanasopoulos, P., Dering, B., Wiggett, A., Kuipers, J. R., & Thierry, G. (2010). Perceptual shift in bilingualism: Brain potentials reveal plasticity in pre-attentive colour perception. *Cognition*, *116*(3), 437–443. https://doi.org/10.1016/j.cognition.2010.05.016

Baus, C., Bas, J., Calabria, M., & Costa, A. (2017). Speak my language and I will remember your face better: An ERP study. *Frontiers in Psychology*, *8*(MAY), 1–12. https://doi.org/10.3389/fpsyg.2017.00709

Bernstein, M. J., Young, S. G., & Hugenberg, K. (2007). The Cross-Category Effect, *18*(8), 706–712.

Blais, C., Jack, R. E., Scheepers, C., Fiset, D., & Caldara, R. (2008). Culture shapes how we look at faces. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0003022

Bodenhausen, G. V., Kang, S. K., & Peery, D. (2012). Social categorization and the perception of social groups. *The SAGE Handbook of Social Cognition*. https://doi.org/http://dx.doi.org/10.4135/9781446247631

Bor, A. (2018). Correcting for base rates in multidimensional "Who said what?" experiments. *Evolution and Human Behavior*, (April), 0–1. https://doi.org/10.1016/j.evolhumbehav.2018.04.003

Bornstein, B. H., Laub, C. E., Meissner, C. a., & Susa, K. J. (2013).

The Cross-Race Effect: Resistant to Instructions. *Journal of Criminology*, *2013*, 1–6. https://doi.org/10.1155/2013/745836
Brewer, M. B., & Yuki, M. (2007). Culture and social identity. In *Handbook of cultural psychology*.
Burns, E., Tree, J., Chan, A., & Xu, H. (2018). Bilingualism shapes the other race effect. *Vision Research*. https://doi.org/10.1016/j.visres.2018.07.004
Chapais, B. (2010). The deep structure of human society: Primate origins and evolution. In *Mind the Gap: Tracing the Origins of Human Universals*. https://doi.org/10.1007/978-3-642-02725-3_2
Chiroro, P. M., Tredoux, C. G., Radaelli, S., & Meissner, C. A. (2008). Recognizing faces across continents: The effect of within-race variations on the own-race bias in face recognition. *Psychonomic Bulletin and Review*. https://doi.org/10.3758/PBR.15.6.1089
Clifford, A., Franklin, A., Holmes, A., Drivonikou, V. G., ??zgen, E., & Davies, I. R. L. (2012). Neural correlates of acquired color category effects. *Brain and Cognition*, *80*(1), 126–143. https://doi.org/10.1016/j.bandc.2012.04.011
Clifford, A., Holmes, A., Davies, I. R. L., & Franklin, A. (2010). Color categories affect pre-attentive color perception. *Biological Psychology*, *85*(2), 275–282. https://doi.org/10.1016/j.biopsycho.2010.07.014
Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences - 2nd Edition. *Hillsdale, NJ: Lawrence Erlbaum Associates*, 8–14. Retrieved from https://www.questia.com/library/98533078/statistical-power-analysis-for-the-behavioral-sciences
Collins, J. A., & Olson, I. R. (2014). Knowledge is power: How conceptual knowledge transforms visual cognition. *Psychonomic Bulletin & Review*, *21*(4), 843–860. https://doi.org/10.3758/s13423-013-0564-3
Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences*, *7*(4), 173–179. https://doi.org/10.1016/S1364-6613(03)00057-3
Crisp, R. J., & Hewstone, M. (1999). Differential evaluation of crossed category groups: Patterns, processes, and reducing intergroup bias. *Group Processes and Intergroup Relations*. https://doi.org/10.1177/1368430299024001
Crisp, R. J., & Hewstone, M. (2007). Multiple Social

Categorization. *Advances in Experimental Social Psychology*. https://doi.org/10.1016/S0065-2601(06)39004-1

Czigler, I. (2014). Visual mismatch negativity and categorization. *Brain Topography*, *27*(4), 590–598. https://doi.org/10.1007/s10548-013-0316-8

Devos, T., & Ma, D. S. (2008). Is Kate Winslet more American than Lucy Liu? The impact of construal processes on the implicit ascription of a national identity. *British Journal of Social Psychology*. https://doi.org/10.1348/014466607X224521

Elliott, E. S., Wills, E. J., & Goldstein, A. G. (1973). The effects of discrimination training on the recognition of white and oriental faces. *Bulletin of the Psychonomic Society*. https://doi.org/10.3758/BF03327717

Firestone, C., & Scholl, B. J. (2014). Can you experience 'top-down' effects on perception?: The case of race categories and perceived lightness. *Psychonomic Bulletin & Review*, *22*(3), 694–700. https://doi.org/10.3758/s13423-014-0711-5

Firestone, C., & Scholl, B. J. (2015). Cognition does not affect perception: Evaluating the evidence for 'top-down' effects. *Behavioral and Brain Sciences*, *4629*(2016), 1–77. https://doi.org/10.1017/S0140525X15000965

Freeman, J. B., & Johnson, K. L. (2016). More Than Meets the Eye: Split-Second Social Perception. *Trends in Cognitive Sciences*, *20*(5), 362–374. https://doi.org/10.1016/j.tics.2016.03.003

Freeman, J. B., Penner, A. M., Saperstein, A., Scheutz, M., & Ambady, N. (2011). Looking the part: social status cues shape race perception. *PloS One*, *6*(9), e25107. https://doi.org/10.1371/journal.pone.0025107

Fuertes, J. N., Gottdiener, W. H., Martin, H., Gilbert, T. C., & Giles, H. (2012). A meta-analysis of the effects of speakers' accents on interpersonal evaluations. *European Journal of Social Psychology*. https://doi.org/10.1002/ejsp.862

Gao, W., Cao, B., Shan, S., Chen, X., Zhou, D., Zhang, X., & Zhao, D. (2008). The CAS-PEAL large-scale chinese face database and baseline evaluations. *IEEE Transactions on Systems, Man, and Cybernetics Part A:Systems and Humans*. https://doi.org/10.1109/TSMCA.2007.909557

Gilbert, C. D., & Li, W. (2013). Top-down influences on visual processing. *Nature Reviews Neuroscience*, *14*(5), 350–363. https://doi.org/10.1038/nrn3476

GILES, H., & Johnson, P. (1987). Ethnolinguistic identity theory: a social psychological approach to language maintenance. *International Journal of the Sociology of Language*, *68*, 69–100. https://doi.org/10.1515/ijsl.1987.68.69

Gramling, O. (1942). *Free Men Are Fighting the story of WWII*. New York, Toronto: Farrar and Rinehart, Inc.

Hansen, K., Rakić, T., & Steffens, M. C. (2017). Competent and warm?: How mismatching appearance and accent influence first impressions. *Experimental Psychology*. https://doi.org/10.1027/1618-3169/a000348

Hehman, E., Stanley, E. M., Gaertner, S. L., & Simons, R. F. (2011). Multiple group membership influences face-recognition: Recall and neurological evidence. *Journal of Experimental Social Psychology*, *47*(6), 1262–1268. https://doi.org/10.1016/j.jesp.2011.05.014

Henrich, J., Heine, S. J., & Norenzayan, A. (2010a). Most people are not WEIRD. *Nature*. https://doi.org/10.1038/466029a

Henrich, J., Heine, S. J., & Norenzayan, A. (2010b). The weirdest people in the world? *Behavioral and Brain Sciences*. https://doi.org/10.1017/S0140525X0999152X

Heron-Delaney, M., Anzures, G., Herbert, J. S., Quinn, P. C., Slater, A. M., Tanaka, J. W., … Pascalis, O. (2011). Perceptual training prevents the emergence of the other race effect during infancy. *PLoS ONE*. https://doi.org/10.1371/journal.pone.0019858

Herrmann, M. J., Schreppel, T., Jäger, D., Koehler, S., Ehlis, a-C., & Fallgatter, a J. (2007). The other-race effect for face perception: an event-related potential study. *Journal of Neural Transmission (Vienna, Austria : 1996)*, *114*(7), 951–7. https://doi.org/10.1007/s00702-007-0624-9

Hewstone, M., Hantzi, A., & Johnston, L. (1991). Social categorization and person memory: The pervasiveness of race as an organizing principle. *European Journal of Social Psychology*. https://doi.org/10.1002/ejsp.2420210606

Hills, P. J. (2012). A developmental study of the own-age face recognition bias in children. *Developmental Psychology*. https://doi.org/10.1037/a0026524

Holmes, A., Franklin, A., Clifford, A., & Davies, I. (2009). Neurophysiological evidence for categorical perception of color. *Brain and Cognition*, *69*(2), 426–434. https://doi.org/10.1016/j.bandc.2008.09.003

Holoien, D. S., & Shelton, J. N. (2012). You deplete me: The cognitive costs of colorblindness on ethnic minorities. *Journal of Experimental Social Psychology*. https://doi.org/10.1016/j.jesp.2011.09.010

Hugenberg, K., Miller, J., & Claypool, H. M. (2007). Categorization and individuation in the cross-race recognition deficit: Toward a solution to an insidious problem. *Journal of Experimental Social Psychology*. https://doi.org/10.1016/j.jesp.2006.02.010

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The Categorization-Individuation Model: An Integrative Account of the Other-Race Recognition Deficit. *Psychological Review*. https://doi.org/10.1037/a0020463

Jackson, M. C., Wilde, V. K., & Goff, P. A. (2016). Seeing color blindness: Color-blind racial ideology research methods in social psychology. In *The myth of racial color blindness: Manifestations, dynamics, and impact*. https://doi.org/10.1037/14754-008

Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: how covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology*, *102*(1), 116–31. https://doi.org/10.1037/a0025335

Kandel, S., Burfin, S., Mï¿½ary, D., Ruiz-Tada, E., Costa, A., & Pascalis, O. (2016). The impact of early bilingualism on face recognition processes. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2016.01080

Kawakami, K., Amodio, D. M., & Hugenberg, K. (2017). *Intergroup Perception and Cognition: An Integrative Framework for Understanding the Causes and Consequences of Social Categorization*. *Advances in Experimental Social Psychology* (1st ed., Vol. 55). Elsevier Inc. https://doi.org/10.1016/bs.aesp.2016.10.001

Kecskés-Kovács, K., Sulykos, I., & Czigler, I. (2013). Is it a face of a woman or a man? Visual mismatch negativity is sensitive to gender category. *Frontiers in Human Neuroscience*, *7*, 1–11. https://doi.org/https://doi.org/10.3389/fnhum.2013.00532

Kenworthy, J. B., Canales, C. J., Weaver, K. D., & Miller, N. (2003). Negative incidental affect and mood congruency in crossed categorization. *Journal of Experimental Social Psychology*. https://doi.org/10.1016/S0022-1031(03)00022-2

Kim, J., & Davis, C. (2010). Knowing what to look for: Voice

affects face race judgements. *Visual Cognition*, *18*(7), 1017–1033. https://doi.org/10.1080/13506281003616057
Kinzler, K. D., Corriveau, K. H., & Harris, P. L. (2011). Children's selective trust in native-accented speakers. *Developmental Science*, *14*(1), 106–11. https://doi.org/10.1111/j.1467-7687.2010.00965.x
Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(30), 12577–80. https://doi.org/10.1073/pnas.0705345104
Kinzler, K. D., Shutts, K., & Correll, J. (2010). Priorities in social categories. *European Journal of Social Psychology*. https://doi.org/10.1002/ejsp.739
Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent Trumps Race in Guiding Children's Social Preferences. *Social Cognition*. https://doi.org/10.1521/soco.2009.27.4.623
Klauer, K. C., Hölzenbein, F., Calanchini, J., & Sherman, J. W. (2014). How malleable is categorization by race? Evidence for competitive category use in social categorization. *Journal of Personality and Social Psychology*, *107*(1), 21–40. https://doi.org/10.1037/a0036609
Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the "who said what?" paradigm. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.75.5.1155
Kovács-Bálint, Z., Stefanics, G., Trunk, a., & Hernádi, I. (2014). Automatic detection of trustworthiness of the face: A visual mismatch negativity study. *Acta Biologica Hungarica*, *65*(1), 1–12. https://doi.org/10.1556/ABiol.65.2014.1.1
Kreegipuu, K., Kuldkepp, N., Sibolt, O., Toom, M., Allik, J., & Näätänen, R. (2013). vMMN for schematic faces: automatic detection of change in emotional expression. *Frontiers in Human Neuroscience*, *7*(October), 714. https://doi.org/10.3389/fnhum.2013.00714
Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(26), 15387–92. https://doi.org/10.1073/pnas.251541498
Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *Journal of*

Abnormal and Social Psychology. https://doi.org/10.1037/h0044430

Lebrecht, S., Pierce, L. J., Tarr, M. J., & Tanaka, J. W. (2009). Perceptual other-race training reduces implicit racial bias. *PloS ONE*, *4*(1), e4215. https://doi.org/10.1371/journal.pone.0004215

Lev-Ari, S. &, & Keysar, B. (2010). Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology*. https://doi.org/10.1016/j.jesp.2010.05.025

Li, X., Lu, Y., Sun, G., Gao, L., & Zhao, L. (2012). Visual mismatch negativity elicited by facial expressions: new evidence from the equiprobable paradigm. *Behavioral and Brain Functions : BBF*, *8*(1), 7. https://doi.org/10.1186/1744-9081-8-7

Liberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). The Origins of Social Categorization. *Trends in Cognitive Sciences*, *xx*, 1–13. https://doi.org/10.1016/j.tics.2017.04.004

Lindsay, D. S., Jack, P. C., & Christian, M. A. (1991). Other-Race Face Perception. *Journal of Applied Psychology*. https://doi.org/10.1037/0021-9010.76.4.587

Liu, A. H. (2017). Democracy and minority language recognition: tyranny of the majority and the conditional effects of group size. *Democratization*. https://doi.org/10.1080/13510347.2016.1245292

Macrae, C. N., & Bodenhausen, G. V. (2000). Social Cognition: Thinking Categorically about Others. *Annual Review of Psychology*. https://doi.org/10.1146/annurev.psych.51.1.93

Malpass, R. S., Lavigueur, H., & Weldon, D. E. (1973). Verbal and visual training in face recognition. *Perception & Psychophysics*. https://doi.org/10.3758/BF03212392

Meissner, C. a., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law*, *7*(1), 3–35. https://doi.org/10.1037//1076-8971.7.1.3

Mepham, K. D., & Martinovic, B. (2018). Multilingualism and Out-Group Acceptance: The Mediating Roles of Cognitive Flexibility and Deprovincialization. *Journal of Language and Social Psychology*. https://doi.org/10.1177/0261927X17706944

Mo, L., Xu, G., Kay, P., & Tan, L.-H. (2011). Electrophysiological evidence for the left-lateralized effect of language on preattentive categorical perception of color. *Proceedings of the National Academy of Sciences*, *108*(34), 14026–14030. https://doi.org/10.1073/pnas.1111860108

Nettle, D. (1999). *Linguistic Diversity*. New York: Oxford University Press.

Neville, H. A., Awad, G. H., Brooks, J. E., Flores, M. P., & Bluemel, J. (2013). Color-blind racial ideology theory, training, and measurement implications in psychology. *American Psychologist*. https://doi.org/10.1037/a0033282

Ng, A. H., Steele, J. R., & Sasaki, J. Y. (2016). Will you remember me? Cultural differences in own-group face recognition biases. *Journal of Experimental Social Psychology*, *64*, 21–26. https://doi.org/10.1016/j.jesp.2016.01.003

Nier, J. A., Gaertner, S. L., Dovidio, J. F., Banker, B. S., Ward, C. M., & Rust, M. C. (2001). Changing Interracial Evaluations and Behavior: The Effects of a Common Group Identity. *Group Processes & Intergroup Relations*. https://doi.org/10.1177/1368430201004004001

Norton, M. I., Sommers, S. R., Apfelbaum, E. P., Pura, N., & Ariely, D. (2006). Color blindness and interracial interaction: Playing the Political Correctness Game. *Psychological Science*. https://doi.org/10.1111/j.1467-9280.2006.01810.x

Noymer, A., Penner, A. M., & Saperstein, A. (2011). Cause of death affects racial classification on death certificates. *PloS One*, *6*(1), e15812. https://doi.org/10.1371/journal.pone.0015812

Palmer, M. A., Brewer, N., & Horry, R. (2013). Understanding gender bias in face recognition: Effects of divided attention at encoding. *Acta Psychologica*. https://doi.org/10.1016/j.actpsy.2013.01.009

Penner, A. M. (2014). The Race of a Criminal Record : How Incarceration Colors Racial Perceptions, *57*(1), 92–113. https://doi.org/10.1525/sp.2010.57.1.92.SP5701

Penner, A. M., & Saperstein, A. (2008). How social status shapes race. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(50), 19628–30. https://doi.org/10.1073/pnas.0805762105

Pietraszewski, D. (2016). Priming Race: Does the Mind Inhibit Categorization by Race at Encoding or Recall? . *Social Psychological and Personality Science* , *7*(1), 85–91.

https://doi.org/10.1177/1948550615602934

Pietraszewski, D., Cosmides, L., & Tooby, J. (2014). The content of our cooperation, not the color of our skin: An alliance detection system regulates categorization by coalition and race, but not sex. *PLoS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0088534

Pietraszewski, D., & Schwartz, A. (2014a). Evidence that accent is a dedicated dimension of social categorization, not a byproduct of coalitional categorization. *Evolution and Human Behavior*, *35*(1), 51–57. https://doi.org/10.1016/j.evolhumbehav.2013.09.005

Pietraszewski, D., & Schwartz, A. (2014b). Evidence that accent is a dimension of social categorization, not a byproduct of perceptual salience, familiarity, or ease-of-processing. *Evolution and Human Behavior*, *35*(1), 43–50. https://doi.org/10.1016/j.evolhumbehav.2013.09.006

Plaut, V. C., Thomas, K. M., & Goren, M. J. (2009). Is multiculturalism or colorblindness better for minorities? *Psychological Science*. https://doi.org/10.1111/j.1467-9280.2009.02318.x

Quinn, P. C., Lee, K., & Pascalis, O. (2018). Perception of Face Race by Infants: Five Developmental Changes. *Child Development Perspectives*. https://doi.org/10.1111/cdep.12286

Rakić, T., Steffens, M. C., & Mummendey, A. (2011). Blinded by the accent! The minor role of looks in ethnic categorization. *Journal of Personality and Social Psychology*, *100*(1), 16–29. https://doi.org/10.1037/a0021522

Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin*. https://doi.org/10.1037/a0025750

Richeson, J. A., & Nussbaum, R. J. (2004). The impact of multiculturalism versus color-blindness on racial bias. *Journal of Experimental Social Psychology*. https://doi.org/10.1016/j.jesp.2003.09.002

Sangrigoli, S., Pallier, C., Argenti, A.-M., Ventureyra, V. A. G., & De Schonen ’, S. (2005). Reversibility of the Other-Race Effect in Face Recognition During Childhood. *Psychological Science*. https://doi.org/10.1111/j.0956-7976.2005.01554.x

Schneider, W., Eschman, a, & Zuccolotto, a. (2002). E-Prime reference guide. *Psychology Software Tools*. https://doi.org/10.1186/1756-0381-3-1

Shriver, E. R., Young, S. G., Hugenberg, K., Bernstein, M. J., & Lanter, J. R. (2008). Class, race, and the face: social context modulates the cross-race effect in face recognition. *Personality & Social Psychology Bulletin*, *34*(2), 260–74. https://doi.org/10.1177/0146167207310455

Sporer, S. L. (2001). Recognizing Faces of Other Ethnic Groups: An Integration of Theories. *Psychology, Public Policy, and Law*. https://doi.org/10.1037/1076-8971.7.1.36

Stangor, C., Lynch, L., Duan, C., & Glass, B. (1992). Categorization of Individuals on the Basis of Multiple Social Features. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.62.2.207

Stefanics, G., Csukly, G., Komlósi, S., Czobor, P., & Czigler, I. (2012). Processing of unattended facial emotions: A visual mismatch negativity study. *NeuroImage*, *59*(3), 3042–3049. https://doi.org/10.1016/j.neuroimage.2011.10.041

Stevenage, S. V., Clarke, G., & McNeill, A. (2012). The "other-accent" effect in voice recognition. *Journal of Cognitive Psychology*, *24*(6), 647–653. https://doi.org/10.1080/20445911.2012.675321

Tanaka, J. W., & Pierce, L. J. (2009). The neural plasticity of other-race face recognition. *Cognitive, Affective & Behavioral Neuroscience*, *9*(1), 122–31. https://doi.org/10.3758/CABN.9.1.122

Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology*, *36*(7), 778–793. https://doi.org/10.1037//0022-3514.36.7.778

Thierry, G., Athanasopoulos, P., Wiggett, A., Dering, B., & Kuipers, J.-R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(11), 4567–70. https://doi.org/10.1073/pnas.0811155106

United Nations. (2017). *International Migration Report 2017. Department of Economic and Social Affairs, United Nations*. https://doi.org/ST/ESA/SER.A/404

Vetter, P., & Newen, A. (2014). Varieties of cognitive penetration in visual perception. *Consciousness and Cognition*, *27*(1), 62–75. https://doi.org/10.1016/j.concog.2014.04.007

Wiese, H. (2012). The role of age and ethnic group in face recognition memory: ERP evidence from a combined own-age and own-race bias study. *Biological Psychology*, *89*(1), 137–147. https://doi.org/10.1016/j.biopsycho.2011.10.002

Wolsko, C., Park, B., Judd, C. M., & Wittenbrink, B. (2000). Framing interethnic ideology: Effects of multicultural and color-blind perspectives on judgments of groups and individuals. *Journal of Personality and Social Psychology*. https://doi.org/10.1037/0022-3514.78.4.635

Wright, D. B., & Sladden, B. (2003). An own gender bias and the importance of hair in face recognition. *Acta Psychologica*. https://doi.org/10.1016/S0001-6918(03)00052-0

Young, S. G., Hugenberg, K., Bernstein, M. J., & Sacco, D. F. (2012). Perception and Motivation in Face Recognition: A Critical Review of Theories of the Cross-Race Effect. *Personality and Social Psychology Review*. https://doi.org/10.1177/1088868311418987

Yu, M., Li, Y., Mo, C., & Mo, L. (2017). Newly learned categories induce pre-attentive categorical perception of faces. *Scientific Reports*. https://doi.org/10.1038/s41598-017-14104-6

Yuki, M. (2003). Intergroup Comparison versus Intragroup Relationships: A Cross-Cultural Examination of Social Identity Theory in North American and East Asian Cultural Contexts. *Social Psychology Quarterly*. https://doi.org/10.2307/1519846

Yuki, M. (2011). Intragroup relationships and intergroup comparisons as two sources of group-based collectivism. In *Social Cognition Social Identity and Intergroup Relations*. https://doi.org/10.4324/9780203816790

Zhao, L., & Li, J. (2006). Visual mismatch negativity elicited by facial expressions under non-attentional condition. *Neuroscience Letters*, *410*(2), 126–131. https://doi.org/10.1016/j.neulet.2006.09.081

Appendix

Sentences used in Chapter 2 and Chapter 3:

	Spanish	English
1	La botella de vino se acaba.	The bottle of wine is finished.
2	La bombilla se ha fundido.	The bulb has gone out.
3	El despertador suena a las dos.	The alarm rings at two.
4	El collar es de diamantes.	The necklace is made of diamonds.
5	El avión despegó sin tripulantes.	The airplane took off without the crew.
6	Dan los dibujos por la tele.	They show cartoons on TV.
7	El suelo está mojado.	The ground is wet.
8	Hoy abrirán las tiendas.	The stores will open today.
9	El coche es muy espacioso.	The car is very roomy.
10	La libreta es lila.	The book is lilac.
11	El móvil se cayó al suelo.	The phone fell to the floor.
12	El casco es rojo y negro.	The helmet is red and black.
13	El diccionario es grueso.	The dictionary is thick
14	El cielo está nublado.	The sky is cloudy.
15	El baúl es muy antiguo.	The trunk is very old.
16	El vestido es muy elegante.	The dress is very elegant.
17	El gimnasio está vacío.	The gym is empty.
18	La almohada es cómoda.	The pillow is comfortable.
19	La camisa está muy sucia.	The shirt is very dirty.
20	El melón tarda en madurar.	The melon takes time to ripen.
21	La casa tiene dos plantas.	The house has two floors.
22	La fiesta ha sido divertida.	The party was fun.
23	El libro tiene cien páginas.	The book has a hundred pages.
24	El ordenador es muy caro.	The computer is very expensive.