

Université de Fribourg
Faculté des Lettres
Département de Psychologie

**Language shifts comprehension processes:
The representation of gender during on-line
reading comprehension among bilinguals of English,
French and German**

Sayaka Sato

Thèse de Doctorat présentée devant la Faculté des Lettres
de l'Université de Fribourg, en Suisse

Sous la direction de Prof. Dr. Raphael Berthele et P. D. Dr. Pascal Gygax

Approuvée par la Faculté des Lettres sur proposition de Prof. Dr. Raphael Berthele (premier rapporteur) et Prof. Dr. Lisa von Stockhausen (deuxième rapporteur).

Fribourg, le 13 janvier 2015
Le Doyen Prof. Dr. Marc-Henry Soulet

Table of Contents

Summary	v
Acknowledgements	vii
List of Tables	ix
List of Figures	ix
List of Abbreviations	xi
1. Introduction	1
2. Theoretical Background	9
2.1 Stereotypes and category labeling	9
2.2 Gender information and language comprehension	12
2.3 The influences of grammatical gender information of person references during gender representation	15
2.4 Grammatical gender and associations through surface forms.....	19
2.5 The linguistic relativity hypothesis and the thinking-for speaking hypothesis	21
2.6 Testing the influences of language among bilinguals.....	25
2.7 Task performance and L2 proficiency	28
2.8 Evaluation of L2 proficiency.....	30
2.9 The bilingual influence on gender representation.....	32
3. Gender inferences: Grammatical features and their impact on the representation of gender in bilinguals	33
3.1 Abstract	33
3.2 Introduction.....	34
3.3 Method.....	42
3.3.1 Participants	42
3.3.2 Materials and design.....	43
3.3.3 Apparatus.....	45
3.3.4 Procedure	46

3.4 Results	46
3.4.1 Proficiency measures.....	48
3.4.2 L1 French speakers	48
3.4.3 L1 English speakers	51
3.5 Discussion and Conclusion	53
4. Representing the social face of grammatical and stereotypical gender:	
When gender-associated stereotypical information does count	59
4.1 Abstract	59
4.2 Introduction.....	60
4.2.1 The impact of language on the construction of gender	
representations	61
4.2.2 The impact of grammaticization of information on mental	
representations across languages	64
4.2.3 The present study	66
4.3 Method.....	68
4.3.1 Participants	68
4.3.2 Materials.....	69
4.3.3 Design and procedure	70
4.3.4 Apparatus.....	71
4.4 Results	72
4.4.1 Proportion of Positive Responses.....	72
4.4.2 Response Times	75
4.5 Discussion	78
4.6 Conclusion	83
5. Women do not operate on patients, men do: Activation of gender	
information during subtle gender occupational descriptions.....	85
5.1 Abstract	85
5.2 Introduction.....	86
5.3 Experiment 1A	92
5.3.1 Method.....	92
5.3.2 Results and discussion	96
5.4 Experiment 1B	104
5.4.1 Method.....	104
5.4.2 Results and discussion	105

5.5 Experimental Comparisons.....	108
5.5.1 L1 French data	109
5.5.2 L2 English data	110
5.6 General discussion	112
5.6.1 L1 French results	112
5.6.2 L2 English results	116
5.7 Conclusion	118
6. Altering male-dominant representations: A study on nominalized adjectives and participles in first and second language German	119
6.1 Abstract	119
6.2 Introduction.....	119
6.2.1 Grammatical gender languages and gender representation	121
6.2.2 Nominalized forms as gender-neutral language in German	123
6.2.3 Cross-linguistic issues on nominalized forms and gender representation	124
6.2.4 The present study	125
6.3 Method.....	126
6.3.1 Participants	126
6.3.2 Design and Materials	127
6.3.3 Apparatus.....	130
6.3.4 Procedure	130
6.4 Results	131
6.4.1 Proportion of positive responses	132
6.4.2 Response times	134
6.5 Discussion and conclusion	136
7. General Discussion	141
7.1 Introduction.....	141
7.2 Summary of the main findings	142
7.3 Theoretical and applied implications	145
7.3.1 Implication of findings for L1 gender representation	145
7.3.2 Implication of findings for L2 gender representation	148
7.4 Limitations and directions for future research	150
7.4.1 Means to address linguistic relativity	150
7.4.2 Investigation of individual differences.....	153

7.5 Conclusion	154
References.....	157
Appendices	173

SUMMARY

The differences in how languages encode gender information may lead to distinctive gender biases when reading and comprehending text. Some languages (e.g., French and German) consist of gender as grammatical and conceptual features while others (e.g., English) only convey gender information conceptually. This thesis investigated this issue with a particular focus on how bilingual¹ readers of English, French and German mentally represent gender when reading person reference role nouns (e.g., *nurses*). Past studies in the L1 (first language) have demonstrated that readers rely on different sources of information during the representation of gender, with English readers basing their representation on their knowledge of gender stereotypes, whereas French and German readers relying more on grammatical gender information than stereotypical information. Based on these differences, this thesis addressed whether these gender biases inherent among the languages may shift in accordance to the language switch bilinguals undergo.

The four sets of studies presented in this thesis provided strong evidence to suggest that the language in which information was being encoded served to guide readers to attend to different types of gender information. Reading a role noun like *nurses* activated stereotype consistent inferences (i.e., nurses = women) in L1 English, whereas the masculine grammatical marking (though interpretable as a generic form) on the role noun in L1 French and L1 German guided readers to represent a male-biased representation even if assumptions about their gender stereotypes countered the interpretation (i.e., *infirmiers*_{Masculine} / *Krankenpfleger*_{Masculine} [= male nurses]). Our findings suggested that the role noun cues incorporate information about the gender of the person and their associated descriptions as part of the lexical representation, which became co-activated during reading processes (Study III). Furthermore, it was claimed that the recurrent male biases observed in L1 French and German were robust, being resistant to the reversing female gender biases brought about by other grammatical forms such as the gendered determiner *die* in German (Study II). In fact, the male bias, at least in German, could only be eliminated when masculine forms were replaced with (gender-neutralizing) *nominalized* forms (Study IV).

¹ The concept of *bilinguals* will be further defined and elaborated in the Theoretical Background section (Chapter 2).

² The term *mental model* in the context of this thesis is used interchangeably with the term *mental representation(s)* and *representation(s)*, and refers to a conceptual entity of symbols and images that reflect

However, shifting to the L2 (second language) was found to alter readers' biases, whereby they switched their representation tendency to conform to the language-bound regularities of the L2. Second language proficiency was found to influence the impact in which the L2-consistent characteristics would manifest (Studies I and II). High L2 proficiency was associated to a gender representation tendency resembling that of native speakers of the bilinguals' L2, whereas low L2 proficiency revealed readers' greater reliance for the L1, which was essentially carried over when processing L2. Importantly, these representational shifts were found to be more difficult if the specific linguistic property was not existent in readers' L1 (Studies I and IV). Moreover, in cases where the bilinguals' linguistic repertory consisted of two grammatical gender languages that brought about a male bias, the L1 was found to trigger a greater bias than the L2 (Studies II and IV).

With the studies providing evidence to suggest that gender representation tendencies were influenced and altered in accordance with the readers' shift in language, these studies shed light on how language directly influences cognitive processes associated to the on-line construction of mental representations.

ACKNOWLEDGEMENTS

The last four years was a period of discovery: a new culture, a new language and new research discoveries. There are no words that would suffice to express my gratitude to the many people who have supported me throughout this endeavor, and it goes without mention that this thesis is the product of their kindness and contribution. I extend my sincere appreciation to all those who made this thesis possible.

First and foremost, I owe my deepest appreciation to my supervisors, *Professor Raphael Berthele* and *Dr. Pascal Gygax* who made my arrival and studies possible at the University of Fribourg. I am grateful to have had the opportunity to work with both of them.

Pascal has been the guiding inspiration throughout these four past years. I cannot express enough my sincere appreciation for his continuous encouragement, guidance, patience and friendship throughout this adventurous road. I benefited from the opportunity to work with a person who was as curious and passionate for research as he was. Both a mentor and a friend, he was truly the supervisor that all Ph.D. students dream for but actually never get. I look forward to more collaboration discussions on the train and lunchtime champagne popping after publications in the future.

I am also indebted to *Professor Ute Gabriel* at the Norwegian University of Science and Technology who is also the co-author of three of the submitted articles included within this thesis. She has been generous with her time and involvement throughout the different stages of this thesis and has treated me as one of her own students.

I would also like to extend my special appreciation to *Professor Yasunori Morishima* at the International Christian University, Japan, for bringing me first into the world of research and encouraging me to pursue my interests and goals. I would not have had the courage to fly all the way to Switzerland alone had you not pushed me to do so.

Furthermore, my gratitude extends to all my colleagues in the Psychology Department at the University of Fribourg especially past and current members of the PPSA *Christelle Gillioz*, *Phillipe Hanchin*, *Pascal Wagner-Egger*, *Arik Lévy*, *Bénédicte Marfaing* and *Julia Misersky*. Our équipe was extraordinary which was a source of wonderful friendships and profound research discussions. I loved coming in to work just to see you all. Special thanks to *Christelle* for her patience and for making my research experience and integration in

Switzerland possible in the most impossible situations. I would have absolutely given up had you not been there.

My sincere appreciation is due to *Professor Alan Garnham* and *Professor Jane Oakhill* for making my research experience possible at the University of Sussex, U.K., and for their expertise in the field.

A heartfelt thanks also extends to *Ilka Wolter*, *Oriane Sarrasin* and *Eimear Finnegan* for all our discussions over after-work drinks in trying to get through our Ph.D. and post-doc lives at Sussex. A special thanks to *Eimear* for her last minute proofreading! Thank goodness you were all there!

I would also like to extend my heartfelt appreciation to *Professor Sabine Sczesny* and *Professor Lisa von Stockhausen*, coordinators of the Marie Curie Initial Training Network “Language, Cognition and Gender” (ITN LCG), as well as all associated fellows and partners of the program. I imagine that most doctoral students are left at a loss when they begin their Ph.D. programs but I have been so fortunate to be part of such a rich network of people, not to mention an amazing support system. I cherish the memories we shared together and hope that our paths will cross in the future.

To *Dr. Shigeyuki Tahara* and my uncle *Dr. Masaharu Sato*, thank you for giving me a second chance. Life is full of ups and downs but in the end, the turns that I took seemed somehow necessary.

I sincerely acknowledge the funding by the European Commission’s Marie Curie ITN LCG project and the Swiss National Science Foundation’s Early Mobility Fund for Early Researchers who made this research possible.

Lastly, I wish to extend my love, respect and gratitude to my family. For my parents and siblings for always believing in me. For my family in Ticino who gave me a place to call home here in far-away Switzerland. And for *Christian*, who has battled the difficult with me at all times on my right shoulder. Thank you.

LIST OF TABLES

Table 1.1	Role nouns chosen from Gabriel et al. (2008) used for the study adapted by Gygax et al. (2008), along with the proportion of men evaluated by each native language group.....	5
Table 4.1	Fixed effects of the reaction time model of the German group with participants and items as random factors, and by-participant and by-item random slopes for task language.	76
Table 4.2	Fixed effects of the reaction time model of the German group with participants and items as random factors, and by-participant and by-item random slopes for task language.	77
Table 6.1	Non-stereotypical role nouns chosen for experimental items.....	129

LIST OF FIGURES

Figure 2.2	The Revised Hierarchical Model (Kroll & Stewart, 1994; Kroll & Tokowicz, 2005)	26
Figure 3.1	Proportion of positive responses of each native group in their respective L1 (error bars indicate standard error of the mean).	49
Figure 3.2	L1 French speakers' proportion of positive responses of each continuation, only in the female stereotype condition. Proficiency is divided into intermediate and advanced groups (error bars indicate standard error of the mean).	50
Figure 3.3	L1 English speakers' proportion of positive responses of each continuation, only in the female stereotype condition. Proficiency is divided into intermediate and advanced groups (error bars indicate standard error of the mean).	52
Figure 4.1	Proportion of positive responses to accept facial images for each native group in each stereotype condition (independent of task language). Error bars indicate standard errors.	73
Figure 5.1	Effects of Pronoun at each level of Stereotype emerging in the pronoun region for high L2 readers. Error bars indicate standard errors.....	102

Figure 5.2	Effects of Pronoun at each level of Stereotype emerging in the pronoun region for low L2 readers. Error bars indicate standard errors.	103
Figure 5.3	Effects of Stereotype at each level of Pronoun at the occupational description region in L1 French. Error bars indicate standard errors.....	106
Figure 5.4	The effect of pronoun observed in L1 French according to each stereotype condition in both experiments. Error bars indicate standard errors.....	110
Figure 5.5	Stereotype effects for each L2 Experiment reflected in first fixation durations. Error bars indicate standard errors.....	111
Figure 6.1	Proportion of positive responses for the native and the non-native German group in each grammatical gender form. Error bars indicate standard errors.....	133

LIST OF ABBREVIATIONS

ANOVA- Analysis of Variance

L1 – First language

L2 – Second language

ms - Milliseconds

SD - Standard deviation

SE – Standard error

FPRT – First pass reading time

FFD – First fixation duration

M – Mean

TRT – Total reading time

RPD – Regression path duration

Chapter 1

Introduction

The processes involved in reading comprehension require readers to represent their interpretation of the textual situation as a mental model² (Johnson-Laird, 1983; Kintsch & van Dijk, 1978; van Dijk & Kintsch, 1983). As readers proceed through the text, information that is explicitly denoted is integrated with information *inferred* from the readers' world knowledge and updated into their representation in working memory (Graesser, Singer, & Trabasso, 1994; Haviland & Clark, 1974). Researchers have paid considerable attention to how certain types of inferences are evoked during these comprehension processes, as they are key to attaining a coherent and integrated mental model. However, the investigation of processes involved in constructing these mental models is complex, as these processes need to take into consideration the various textual elements that may come into effect.

Among the various interactions that can be observed during reading processes, this thesis focuses specifically on the interplay of grammatical information and conceptual information. While the former signifies the formal linguistic functions that govern language processing, the latter, as defined by Barsalou (1993), refers to people's broad knowledge source of categorical information that may impact cognitive processes such as language comprehension. The challenges resulting from the collective effects of the two sources of information are illustrated in the following classic riddle:

"A man and his son were away for a trip. They were driving along the highway when they had a terrible accident. The man was killed outright but the son was alive, although badly injured. The son was rushed to the hospital and was to have an emergency operation. On entering the operating theatre, the surgeon looked at the boy, and said, "I can't do this operation. This boy is my son". The question is, how can this be?" (Sanford, 1985, p. 311)

² The term *mental model* in the context of this thesis is used interchangeably with the term *mental representation(s)* and *representation(s)*, and refers to a conceptual entity of symbols and images that reflect reality or information entities that have been processed by the perceiver or reader.

In order to understand this riddle, readers must first decode general surface level information. For instance, the words *man* and *son* semantically denote the male gender of the two characters, as well as their familial relation. However, problems can arise if readers generate particular gender-associated inferences about the depicted surgeon based on assumptions of their world knowledge. Specifically, if reading *surgeon* activates a male gender (given that general stereotypical expectations may encourage readers to do so), the incongruity of information that follows (i.e., the surgeon cannot be a father because he has died, hence the surgeon must be the mother) would impede comprehension.

The present thesis deals with this type of conceptual knowledge about a person's gender, also known as *gender stereotypical* information. Gender stereotypical information in the context of the following studies not only refers to the information regarding the perceived frequency of the female and male divide in society (e.g., there are typically more male than female surgeons), but also incorporates an aspect of perceivers' social expectations. For example, the belief that men make more competent surgeons or that women would make for better nurses given that they are good caregivers, regardless of the true validity of these social assumptions.

However, making inferences about the gender of a person is not solely dependent on readers' world knowledge. A similar text presented in a grammatical gender language like French or German where gender information is *formally* marked on a grammatical level may pose different comprehension difficulties. This is due to the fact that there is also explicit grammatical gender information that readers need to take into account when reading. For instance in Sanford's example illustrated above, the word *surgeon* would be marked for the specific gender of the person being referred to, with the feminine marking to indicate a woman (FRENCH: *la chirurgienne*_{Feminine}, GERMAN: *die Chirurgin*_{Feminine} [the female surgeon]) or the masculine form (FRENCH: *le chirurgien*_{Masculine}, GERMAN: *der Chirurg*_{Masculine} [the male surgeon]) to indicate a man, while readers may also infer an implicit stereotypical gender simultaneously.

This thesis aims to further contribute to the understanding of how readers make use of these different types of gender information in different languages during on-line³ reading processes. It focuses particularly on the differences in gender representation in natural gender languages such as English where there is no formal grammatical gender information, and in

³ While *on-line* processing refers to readers' real time processing *during* reading, *off-line* processing refers to post-reading modes that include readers' reflection and interpretation of what they processed during reading.

grammatical gender languages such as French and German where gender information is also embedded as a grammatical feature. Indeed, differences in gender representation between these two language typologies have recently been outlined by Gygax, Gabriel, Sarasin, Oakhill, and Garnham (2008). The authors have suggested that when formal grammatical information is lacking (e.g., languages such as English), gender representation is based on stereotype information, whereas when gender information is also embedded as a grammatical feature (e.g., languages such as French and German), grammatical information plays a more deterministic role in the representation process.

To the extent that languages differ as to how they represent the sexes, examining the interaction of the two sources of gender information makes gender representation a convenient workbench on which to illustrate the integrative processes during the construction of mental models in general. In particular, issues pertaining to how languages may influence our representations and thought processes are addressed throughout the studies, provided that gender is represented differently according to each language typology (Corbett, 1991).

Importantly, the thesis focuses on a population that spoke more than one language (i.e., hereafter bilinguals - our definition will be provided in the Theoretical Background section in Chapter 2). It works under the premise that there are different language-bound constraints in reading comprehension processes and as such, gender biases were expected to shift when the comprehension language switched from one language to another. In other words, if English readers make inferences about gender based on gender stereotypical knowledge, whereas French and German readers rely more on the grammatical rules governing the language, it is reasonable to assume that readers who are able to process in two typologically different languages (i.e., bilinguals) would show different comprehension processes according to each language. Alternatively, it is also possible that bilinguals may present comprehension tendencies that do not represent the characteristics of either language, but rather a tendency that combines (i.e., in-between) the characteristics of the two languages. Ultimately, it is aimed to examine whether a shift in comprehension language would alter readers' representations according to the language in which comprehension is taking place and whether certain biases consistently emerge in each corresponding language. The investigation of such effects should essentially shed light as to whether mental representations in the less dominant second language (L2) can influence readers' perceptive biases, and more importantly, clarify the relationship of the influences of language on cognitive processes.

In what follows, the reported collection of studies⁴ aims to deconstruct the elements that contribute to understanding the relationship between language and mental representations, specifically focusing on grammatical gender and gender stereotypical information. The studies presented in the following chapters take English, French and German as comparative languages for the following reasons: First, English is a language that is not systematically grammatically marked for gender as in French and German (except for certain nouns that semantically denote a specific sex [e.g., *postman*, *policeman*] and pronouns [e.g., *she*]), therefore readers' stereotypical knowledge should or may have a primary effect on gender representation. As for the French and German, they provide exemplars for how grammatical gender and stereotypical information may interact, although the two languages posit different grammatical gender structures.

Study I (Chapter 3) compares gender representation processes between English, a natural gender language with that of French, a grammatical gender language. The two grammatical gender languages, German and French are directly compared in Study II (Chapter 4), where German provides an interesting case for female gender associations on a surface level, as opposed to French, which has no associations of the like. The effects of grammatical and stereotypical cues with different influential magnitudes are examined in Study III (Chapter 5) taking French and English again as comparative languages. And finally in the last study (Study IV, Chapter 6) the neutral nominalization forms in German are tested among native French speakers, given that an equivalent form does not exist in the French language. All experimental studies employed gender-stereotypical occupational role nouns (see Table 1.1 at the last page of the section) in order to examine how stereotypical (i.e., world knowledge) and grammatical information would interact, as has been done in a bulk of psycholinguistic studies on gender representation (e.g., Cacciari, Corradini, Padovani, & Carreiras, 2011; Manuel Carreiras, Garnham, Oakhill, & Cain, 1996; Esaulova, Reali, & von Stockhausen, 2013; Gabriel & Gygax, 2008; Garnham, Gabriel, Sarrasin, Gygax, & Oakhill, 2012; Gygax, Gabriel, Lévy, Pool, Grivel, & Pedrazzini, 2012; Gygax & Gabriel, 2008; Irmen, 2007; Irmen & Kurovskaja, 2010; Irmen & Roßberg, 2004; Lévy, Gygax, & Gabriel, 2014; Reali, Esaulova, & von Stockhausen, 2015; Stahlberg, Sczesny, & Braun, 2001; Steiger-Loerbroks & von Stockhausen, 2014). These role nouns were normed for

⁴ The term *study* will be used to refer to the entire research presented in each experimental chapter, whereas the term *experiment* will be used to refer to the specific empirical experiments conducted in each study herein.

Table 1.1 Role nouns chosen from Gabriel et al. (2008) used for the study adapted by Gygax et al. (2008), along with the proportion of men evaluated by each native language group.

English	%	German	%	French	%
Male stereotypes					
Spies	73	Spione	67	Espions	74
Golfers	73	Golfspieler	68	Golfeurs	73
Politicians	71	Politiker	69	Politiciens	72
Police officers	63	Polizisten	69	Policiers	70
Statisticians	70	Statistiker	72	Statisticiens	74
Bosses	62	Arbeitgeber	72	Patrons	74
Computer specialists	70	Informatiker	79	Informaticiens	67
Surgeons	62	Chirurgen	75	Chirurgiens	75
Technicians	72	Techniker	78	Techniciens	75
Engineers	78	Ingenieure	78	Ingénieurs	74
Physics students	56	Physikstudenten	81	Etudiants en physique	67
Pilots	70	Flieger	76	Aviateurs	74
Mean	68		74		72
Neutral Stereotypes					
Singers	53	Sänger	45	Chanteurs	48
Pedestrians	49	Spaziergänger	46	Promeneurs	52
Cinema goers	51	Kinobesucher	49	Spectateurs de cinéma	50
Concertgoers	47	Konzert-Zuhörer	47	Auditeurs de concert	51
Schoolchildren	53	Schüler	48	Ecoliers	53
Spectators	55	Zuschauer	41	Spectateurs	51
Neighbours	50	Nachbarn	50	Voisins	50
Swimmers	50	Schwimmer	50	Nageurs	50
Tennis players	53	Tennisspieler	52	Joueurs de tennis	54
Authors	48	Autoren	52	Auteurs	54
Musicians	54	Musiker	50	Musiciens	59
Skiers	55	Skifahrer	53	Skieurs	55
Mean	52		49		52
Female stereotypes					
Beauticians	29	Kosmetiker	11	Esthéticiens	18
Birth attendants	29	Geburtshelfer	11	Assistants maternels	18
Fortune tellers	32	Wahrsager	24	Diseurs de bonne aventure	28
Cashiers	39	Kassierer	27	Caissiers	24
Nurses	30	Krankenpfleger	24	Infirmiers	30
Hairdressers	48	Coiffeure	21	Coiffeurs	38
Psychology students	38	Psychologiestudenten	25	Etudiants en psychologie	33
Dieticians	39	Diätberater	27	Diététiciens	37
Dressmakers	43	Schneider/Näher	23	Couturiers	40
Dancers	32	Tänzer	33	Danseurs	29
Sales assistants	34	Verkäufer	33	Vendeurs	37
Social workers	29	Sozialarbeiter	41	Assistants sociaux	33
Mean	35		24		30

stereotypicality in each language, in a norming study by Gabriel, Gygax, Sarasin, Garnham and Oakhill (2008)⁵.

The present thesis begins (Chapter 3 [Study I]) by examining how language shifts in bilinguals may result in a change in gender representation processes, based on the study by Gygax et al. (2008) that revealed typological differences between grammatical and natural gender languages. The results of the reported study further corroborated their studies, but also revealed a representational shift according to the language-bound characteristics. More importantly, an interesting effect of proficiency was revealed that influenced the magnitude of the effect of representational shifts.

Crucially, Study I (Chapter 3) suggested that grammatical gender language readers make the male gender more salient through the association of *surface forms* of masculine forms. In the study presented in Study II (Chapter 4), the influences of these linguistic encodings of gender on surface forms were further examined, in combination with the impact of stereotype information that were overridden in Study I. In particular, following Rothermund (1998), the effects of the female-associated German determiner *sie* was investigated in comparison to the French determiner *les* which has no female or male association. A combined visual-lexical paradigm was implemented to determine whether the locus of gender information activation occurs immediately after reading the stereotypically gendered role noun, and whether these surface associations are influential at conceptual levels of representation. The results suggested that the female association of determiners were superficial, and further confirmed the robustness and strength of the male biases inherent in grammatical gender languages, especially the male bias carried from the L1 to L2.

The view that the activation of gender information is robust was central to Study III (Chapter 5), which investigated whether readers would activate gender information in the presence of more discrete cues. Processing of occupational descriptions was compared to that of role nouns among French-English bilinguals. The study revealed that although less salient cues (i.e., descriptions) associated to stereotypical occupational role nouns were associated to a specific gender, they were not strong enough to immediately activate a specific gender.

Finally, Study IV (Chapter 6) investigated German nominalization forms of adjectives and participles as possible neutralizing forms that activate gender-equal representations. Native German and non-native German bilinguals (French-German) were presented with stereotypically neutral role nouns in the nominalized and masculine forms. The study revealed

⁵ Note that a more recent norming study of stereotypical role nouns has been released by Misersky et al. (2013) that covers Czech, English, French, German, Italian, Norwegian and Slovak.

that although nominalized forms were effective in neutralizing the male bias found when reading the role noun in the masculine form for native German readers, such a neutralizing form did not surface for non-native German speakers.

Together, all presented studies were intended to provide a better understanding of how formal grammatical systems interact with conceptual stereotypical information (which shapes how people represent gender), and whether these biases and representations can be shaped or even changed according to the language in which we process information. They touch on the fact that semantic and conceptual representations are closely related and that language has a significant impact on the construction of these representations. These findings are discussed in further detail in the General Discussion section (Chapter 7) which offers a summary of the findings and draws conclusions with regard to the traditional linguistic relativity hypothesis by Whorf (1956) and the more adapted *thinking-for-speaking hypothesis* by Slobin (1996a, 1996b, 2003). We begin first by clarifying the relevant theoretical background pertaining to the thesis in the following chapter.

Chapter 2

Theoretical Background

2.1 Stereotypes and category labeling

A central issue in social cognitive theories concerns the processes in which people make impressions about other people. Perceivers need to construct coherent mental models (i.e., representations) of the immediate stimuli around them (e.g., Johnson-Laird, 1983; Sherman, 1996; Srull & Wyer Jr., 1983) in order to make sense of the environment and people they encounter. The processes involved during reading comprehension (which this thesis focuses on) are comparable to these theories concerning person perception formation, in that mental models need to be constructed to simulate the information readers understand from text.

Constructing mental models to reflect stimuli however, poses substantial processing demands upon a perceiver's limited cognitive resources given the abundant amount of information that is accessible to them. This task can be facilitated by a strategy known as *social categorization* which increases predictability by allowing perceivers to direct their attention only to relevant information, while ignoring those that are of lesser importance. This notion suggests that during person perception, people are grouped according to their trait and behavioral similarities, predicted from preexisting beliefs about a social group that they belong to (e.g., ethnicity, gender, age).

Also known as *stereotypes*, such knowledge structures function as cognitive schemas (Fiske & Taylor, 1991; Macrae & Bodenhausen, 2000), emphasizing the idea that there are general mental representations containing abstract forms of typicality and beliefs about specific social groups. Applying these knowledge structures subsequently enables associated information to be inferred and applied, facilitating person recognition and impression formation.

Although this basic categorization mechanism is fundamentally convenient that maximizes cognitive efficiency (Allport, 1954; Macrae, Milne, & Bodenhausen, 1994), stereotypical beliefs represent only a prototype of traits and characteristics that do not always result in a suitable representation of reality. In fact, the application of stereotypes assumes that information may become biased, as the incoming information would need to be fitted into

rigid constructs. A common shared belief against a group may be based on false information, and successively lead to the formation of negative (and false) stereotypes of all members of that group. Additionally, stereotype application implies that all members of the group would be perceived as sharing the same traits limiting any leeway for individual characteristic differences (Hilton & von Hippel, 1996). With reference to Sanford's example presented earlier, stereotypes of a prototypical surgeon may incorporate representations that assume they are usually male and have male-like characteristics such as being decisive and intelligent. These assumptions may inevitably influence readers' representations, regardless of the absence of such implications in the text; for instance hindering the possibility that a surgeon could be a woman or a mother.

Recently, the notion of stereotypes has been inspected alongside language, as language is one of the primary means stereotypes are carried and communicated to a wider mass (Maas & Arcuri, 1996). Language reflects the way society views certain social groups, communicating stereotypical assumptions that have direct and concrete social consequences. This view of language promoting stereotypical beliefs can be seen in the practice of category labeling, which supposes that people may be labeled differently according to their stereotypical functions.

For example, although referring to the same group of persons, the terms *nigger* and *African-American* obviously instantiate different images of these people (Maas & Arcuri, 1996). However, while discernible derogatory terms of this sort such as *nigger* explicitly represent a common stigmatized view that a certain community may share, other practices of category labeling that promote implicit stereotypes can be more discrete. For instance, feminist linguists claim that linguistic conventions reflect how men have traditionally occupied major social occupational roles. Gender specifying suffixes such as *-man* seen on high-status role nouns such as "chairman" and "spokesman" endorse the idea that these roles and occupations should specifically be held by men. In contrast, suffixes that denote the female sex (e.g., *-ess*) in occupations such as "stewardess" and "waitress" accentuate the fact that jobs that serve others should consist of women. These gender-specifying suffixes found in occupations further contribute in fostering social expectations of what each gender can or cannot do.

The basic model that can explain the relationship between such linguistic practices of category labeling (i.e., language) and *stereotype activation*, stems from the idea of spreading activation of semantic network models (Collins & Loftus, 1975). This model suggests that nodes of information are lexically and semantically interconnected within long-term memory.

When a single semantic node within this model is activated, its activation spreads within the network, co-activating related nodes and making them more accessible. This line of reasoning assumes that activation of a single categorical label would co-activate associated stereotypes such as related personality traits or behavioral expectations. Consequently, the representation of characteristics that match or are associated in some way (i.e., stereotypical traits) would therefore have a greater advantage in terms of activation facilitation than those that are not connected or are less associated.

This notion has been empirically shown in early research that has demonstrated stereotype influences using priming studies. Largely, these studies have indicated that when primed with a social categorical group such as a specific race, responses to stereotypically associated traits were more facilitated than those that were not associated (e.g., *black - lazy* vs. *white - lazy*, Dovidio, Evans, & Tyler, 1986). Stereotype activation effects were found even when primes were presented subliminally so as to mask the relationship between the prime and target (Devine, 1989; Lepore & Brown, 1997; Macrae et al., 1994).

With regard to gender stereotype information which this thesis is based on, results from semantic priming paradigms using match-mismatch effects parallel the above-mentioned accounts. For example, Banaji and Hardin (1996) presented participants with primes that were gender definitional (e.g., *mother*), gender stereotypical (e.g., *nurse*) or gender neutral (e.g., *student*), that were followed by target pronouns (*she / he*). In the first experiment, participants had to decide whether the target pronoun was male or female (e.g., *he* vs. *she*) and, in the second experiment, they had to judge whether the presented target word was a pronoun or not (e.g., *she* or *it* vs. *do* or *as*). In both experiments, they found that participants' reaction times were faster when there was a gender match between the prime and targets. In a similar vein, using a word-pair judgment paradigm, Oakhill, Garnham, and Reynolds (2005) presented participants with pairs of stereotypical occupational role nouns and gender specific kinship terms (e.g., *surgeon - brother* vs. *surgeon - sister*). When asked to judge whether the pairs could refer to the same person or not, judgment times were faster when there was a gender match between the role noun and the kinship term. Interestingly, they found that suppression of stereotype activation was difficult even when participants were encouraged to do so, revealing that gender stereotypical information was immediately activated and incorporated into the mental representation with effects being persistent and strong.

These studies highlight that the mechanism of stereotype activation operates on an *automatic*⁶ and unintentional level that may extend its influence to subsequent judgments and tasks (Banaji & Greenwald, 1995). In the following section, the influences of gender information are further examined within the context of reading comprehension.

2.2 Gender information and language comprehension

Assuming the general influences of language on social cognition, recent research has focused on investigating the relationship between language and *gender*, as gender is one of the most fundamental and salient dimensions of incoming information people rely on during person perception (Bower & Karlin, 1974; Bruner, 1957; Fiske, 1998).

Conveyed both grammatically and conceptually (Stahlberg, Braun, Irmen, & Sczesny, 2007), gender has a significant impact upon how different language speakers comprehend languages (Corbett, 1991). Three basic language categories, as defined by Stahlberg et al. (2007), illustrate the different extent to which gender is encoded within languages and essentially provide a suitable framework to determine the types of information contributing to gender representation.

The first category known as *natural gender languages* consists of languages such as English and those of the Scandinavian language group. These languages do not inherently carry grammatical gender but words such as person nouns (e.g., *king* and *sister*) and pronouns (e.g., *she* and *he*) may carry an explicit gender that is semantically associated to the actual biological sex of the referent. The expression of gender information can easily be escaped compared to the second category of languages known as grammatical gender languages, given that sex does not necessarily need to be linguistically encoded. The second category, *grammatical gender languages*, consists of languages belonging to Slavic, German, Romance, Indo Aryan or Semitic languages, and is considered to have the highest gender saliency among the three categories. With the exception of person nouns, all nouns within this group are arbitrarily assigned a gender (e.g., FRENCH: *la lune*_{Feminine} [the moon], *le soleil*_{Masculine} [the sun]), whereas person nouns have a grammatical gender matching the biological sex of the referent (FRENCH: *l'infirmière*_{Feminine} [the female nurse]; *l'infirmier*_{Masculine} [the male nurse]). Gender agreement influences all linguistic dependencies such as adjectives and articles that must carry the same gender as its corresponding noun.

⁶ Although there is a substantial debate in psychology as to the usage of the term automaticity (e.g., Bargh, 1999), the term is used throughout the thesis to nuance the notion of information activation being spontaneous.

Expression of gender information is unavoidable given that it is a fundamental grammatical feature. Finally, the third category, *genderless languages* refers to languages belonging to Uralic, Turkic, Iranian, Sinitic and Bantu languages, and are those in which there are no linguistic distinctions between the two genders.

The lack of a consistent and formal grammatical feature of gender in natural gender languages implies that gender representation in language comprehension, especially during reading, may be heavily reliant on language comprehenders' tacit world knowledge. This assumption concurs with the aforementioned priming studies in English by Banaji and Hardin (1996) and Oakhill et al. (2005), that demonstrated an automatic activation of gender information based on stereotypical world knowledge.

Psycholinguistic studies using on-line sentence comprehension paradigms also provide converging results. Kennison and Trofe (2003) investigated the effects of stereotypical information during sentence comprehension. Participants were presented with sentence pairs that included stereotypically female (e.g., *secretary*) or male (e.g., *accountant*) role nouns in the first sentence that were followed by a second sentence including either a female or male specifying pronoun (*she* vs. *he*). The authors found gender mismatch effects reflected in prolonged reading times when there was a gender violation between the gender of the role noun and pronoun. Similarly, in a more recent study by Gygax, Gabriel, Sarasin, Oakhill and Garnham (English data, 2008), English readers were presented with a sentence evaluation task (the French and German data will be discussed later). In the first sentence, role nouns with female (e.g., *nurses*), male (e.g., *surgeons*) or neutral (e.g., *pedestrians*) stereotypes were mentioned, followed by a second sentence that referred back to the role noun group with a continuation expression specifying a group of men or women (i.e., *some of the men / some of the women*). The task that followed required participants to judge the plausibility of whether the second sentence was an acceptable continuation to the first sentence. The results indicated that English readers relied on gender stereotypical knowledge, and accepted continuation sentences with gender references that matched the gender stereotypicality of the preceding role noun (i.e., *nurses* - some of the women; *surgeons* – some of the men).

Models of reading comprehension have clarified the mechanisms underlying the activation of gender stereotype information during text comprehension within the framework of the mental models approach. The account assumes that readers generate stereotype-based inferences based on world knowledge to fill in gaps of unspecified information, and consistently update this information into their mental model (Garnham, 2001). These

inferences are essentially generated to deepen the representation even when it is not textually specified as necessary information.

While these studies highlight that stereotype information is spontaneously activated and mentally represented in natural gender languages, there is an additional grammatical component that adds to the representation process for grammatical gender languages. Indeed, research has identified the impact of grammatical gender information on various linguistic processing mechanisms, including noun recognition (Dahan, Swingley, Tanenhaus, & Magnuson, 2000; Grosjean, Dommergues, Cornu, Guillelmon, & Besson, 1994; Wicha, Moreno, & Kutas, 2003), text comprehension (Cacciari et al., 2011; Carreiras et al., 1996; Desrochers, 1986), production (Vigliocco & Franck, 2001) and even gender attribution (to unfamiliar nouns) for children at very early stages of cognitive development (Karmiloff-Smith, 1979). Nonetheless, the influences of grammatical gender on comprehension processes have been difficult to grasp given that they affect not only formal grammatical features, but also appear to *semantically* contribute to the representation of gender.

Numerous studies have reported that knowledge of grammatical gender information in a language may influence one's perception of objects and animals. For instance, Konishi (1993) investigated the influence of grammatical gender information on the personification of nouns. German and Spanish speaking adult participants (mean age: 25.3 years [women], 20.8 years [men]) were instructed to rate a set of nouns for their potency levels (e.g., weak-strong), a dimension that is closely related to male attributes, on a scale of 1 (e.g., *extremely weak*) to 7 (e.g., *extremely strong*). The results suggested that words that were grammatically masculine in each language were rated higher on the rating scale suggesting a stronger association to the male gender than words that were grammatically feminine. Similarly, Sera, Berge and del Castillo-Pintado (1994) instructed their participants to attribute a female or male voice to images of animals and everyday objects. Unlike their English counterparts, gender attribution for Spanish adult speakers was consistent with the grammatical gender of the object. However, Spanish children's (mean age: 5 years, 4 months) responses were not as consistent, with the authors arguing that this results from the acquisition of grammatical gender properties.

One methodological concern raised by Boroditsky, Schmidt, and Phillips (2003) regarding the above studies was that participants were instructed to make subjective responses, which may have resulted in forming a strategy to respond to the task. To counter these possible problems, the authors made the connection of gender information and the experimental task more covert. They found that when specific proper names were given to

objects (e.g., *Patrick*_{Male} or *Patricia*_{Female} for an apple), Spanish and German speaking participants were able to recall more objects if the gender of the names matched the grammatical gender of the objects in their own language. Furthermore, adjectives describing inanimate objects were associated to gender traits that matched the items' grammatical gender (e.g., *Key*: German [masculine] - hard, heavy, jagged; Spanish [feminine] - beautiful, elegant, fragile). Crucially, assigning a gender to an inanimate object also triggered the production of associated gender traits such as *beautiful* and *hard* (Boroditsky et al., 2003).

In sum, these findings suggest that for languages with grammatical gender, the grammatical gender distinctions serve to distinguish category differences. Perceivers find that entities with the same grammatical gender share similarities. Consequently, people apply meaningful characteristics to these entities, based on gender characteristics typical to the gender category (e.g., masculine – strong) in question. If such information embedded in grammar does not serve purely grammatically categorical functions, but is also semantically informative, the implications this may serve for speakers of grammatical gender language comprehenders should be taken into serious consideration. This is because grammatical gender information is closely tied to the conceptual sex of person references, which may result in impacting reading processes.

2.3 The influences of grammatical gender information of person references during gender representation

When referring to a person in grammatical gender languages, there is an assumption that the sex of the person will be denoted through the use of explicit grammatical forms. The rules for gender allocation to a specific person reference however is relatively intricate. In fact, research has shown that the actual process involved in the representation of gender can be complicated by several impacting factors.

The first issue relates to the interaction between multiple sources of gender information that are activated simultaneously. As previously illustrated, the impact of stereotypical assumptions during reading is crucial; readers activate a specific gender based on general stereotypical expectations when encountering role nouns. Grammatical gender readers however, are also faced with additional grammatical gender information that impact gender representation. For instance, to examine the influence of stereotypical information in grammatical gender languages, Carreiras et al. (1996) presented participants with texts that included stereotypical role nouns (e.g., *doctor*, *nurse*) followed by sentences including a

pronoun (*he /she*) in both English and Spanish. They found prolonged reading times on the second sentence in English, when the pronoun gender mismatched the stereotypical gender of the role noun, indicating a stereotype influence for gender representation. In contrast, no mismatch effects were observed on the second sentence in Spanish given that the grammatical marking in the first sentence had already specified the role noun gender. Importantly however, inflated reading times were observed on the first sentence when the gender stereotypicality of the role noun mismatched the gender specified by the gender marked determiner (*la_{Feminine} futbolista* [female footballer]). Their results revealed an interaction of stereotypical and grammatical information, suggesting that gender representation processes may be hindered when gender information implied by stereotypical information violates the gender explicitly specified by grammatical gender.

While the interaction of stereotype and grammatical gender information seems evident, research examining the time course in which these sources of information interact show mixed results. Lexical based paradigms (as proposed by Banaji and Hardin, 1996 and Oakhill et al., 2005 described earlier) conducted in grammatical gender languages reported immediate activation of stereotype effects when encountering the role noun (e.g., Gygax et al., 2012; Lévy et al., 2014; Cacciari & Padovani, 2007). These results were also qualified by sentence-based studies incorporating eye-tracking that observed early stereotype influences. For instance, Irmen (2007) presented participants with stereotypical role nouns in the masculine form (e.g., *Soldaten* [soldiers]) followed by sentences including gender specifying anaphoric phrases (e.g., these men / these women). Results showed that when stereotypical gender violated the grammatical gender of the role noun, reading time was inflated immediately before and after the anaphoric noun, indicating an early stereotype effect at lexical access. Grammatical gender effects however surfaced only at the anaphor, indicating a later effect.

In contrast, Esaulova, Reali and von Stockhausen (2013) reported an opposite tendency. In a similar eye-tracking study to that of Irmen (2007), participants read sentences including a stereotypical role noun (*Elektriker* [electrician]) followed by anaphors (Experiment 1: pronouns, Experiment 2: noun phrases). Gender match-mismatch effects indicated early effects of grammatical gender information and later effects of stereotype information. The authors attributed the differences in findings of their study to previous studies as being dependent on the exact materials and procedures employed in the conducted studies.

The second issue that contributes to complicating gender representation in grammatical gender languages relates to the interpretation ambiguity of the masculine grammatical form. In most grammatical gender languages, while the feminine form unequivocally refers to a female referent, the masculine form *can* be used (a) to refer to men only, (b) when the gender of the referent is unimportant or unknown, or (c) to refer to a group of both genders rather than using a third grammatical form. This means that the masculine form is used to signify both a male-specific and *generic* referent, thus leading to interpretation ambiguity and a potential reading complication.

Indeed, studies show that there may be qualitative differences in how the masculine and feminine forms are represented. For instance, Irmen and Schumann (2011) presented gender definitional kinship nouns (e.g., brother) followed by masculine or feminine role nouns in German sentences. A slowdown in processing role nouns with feminine grammar (e.g., *Sängerin*_{Feminine} [singer]) was observed at early stages of processing when its kinship gender mismatched (e.g., brother), although this was not the case for masculine grammatical role nouns. The authors argued that these results were indicative of different processing mechanisms for feminine and masculine grammatical forms. While the latter masculine form is ambiguous and open to referential gender, the former feminine form is unambiguous. If there is a gender violation, the feminine form results in an immediate slow down, whereas the masculine form is underspecified in gender and therefore does not immediately exhibit the gender incongruency. Importantly, they concluded that despite the different processing mechanisms, there were no differences in terms of processing difficulty.

In fact, it has recently been argued by feminist linguists that use of the masculine form to denote a generic interpretation works unfavorably for women's social visibility (Braun, Sczesny, & Stahlberg, 2005; Peyer & Wyss, 1998; Stahlberg et al., 2001). In other words, because grammatical gender is related to the conceptual sex of the person being referenced, the interpretation of the masculine form would most likely facilitate a cognitive representation of men. The critique is reasonable, as only one man in a group of women suffices for the masculine form to be applied. In this sense, using the masculine form potentially adds to the association of the male gender to the mental representation (French: Brauer & Landry, 2008; Chatard, Guimont, & Martinot, 2005; German: Braun et al., 2005; Norwegian: Gabriel, 2008; Gabriel & Gygax, 2008).

In support of these claims, on-line and off-line studies in German (Braun, Gottburgsen, Sczesny, & Stahlberg, 1998; Stahlberg, Braun, Irmen, & Sczesny, 2007; Stahlberg & Sczesny, 2001) and French (Brauer & Landry, 2008; Gygax et al., 2012, 2008;

Lévy et al., 2014) have largely suggested that readers are likely to construct representations that are male dominated when presented with masculine forms, even when stereotypical information may support representation of the female gender (e.g., *infirmiers*_{Masculine} [male nurses]).

Indeed, contrary to the English readers who showed a reliance on stereotypical information for gender representation (i.e., gender match effect), the French and German readers in the aforementioned study (section 2.2) by Gygax et al. (2008), were more likely to base their representations on grammatical gender information. Note that the role nouns in French and German were presented in the masculine plural forms that could signify either a generic (referring to both men and women) or a male-specific reference (e.g., French: *infirmiers*_{Masculine plural} [nurses]; German: *Krankenpfleger*_{Masculine plural} [nurses]). The results showed that irrespective of the gender stereotypicality of the role noun, sentence continuations referring to men were judged as being more plausible than those with reference to women when they followed sentences that consisted of the masculine form (i.e., *infirmiers*_{Masculine plural} [nurses] – *plusieurs hommes* [some of the men]). The authors argued that the *specific* meaning of the masculine form overrode stereotype information, resulting in a male-dominant representation.

Likewise, Braun et al. (2005), through a series of off-line experiments, showed that participants were less likely to mentally include women into their representations (e.g., estimate the proportion of women attending a scientific meeting) when they read articles and filled out questionnaires that were written in the masculine form (intended as generic) than when other more gender-fair word forms were used. The authors attributed the preference of the male-specific interpretation of the masculine form to factors such as higher frequency of the male-specific interpretation, general difficulty in identifying which interpretation is being conveyed, and the fact that context effects also add to the specification of the male gender.

Together, these studies attest that formal grammatical gender information outweighs conceptual information for grammatical gender language readers. More importantly, the masculine form (that is theoretically open to multiple interpretations) was found to further emphasize the male gender through association of its *surface forms* with grammatically masculine forms, influencing cognitive representations. As these examples show, surface forms of grammatical gender appear to play an ‘emphasizing’ role in the gender representation process.

2.4 Grammatical gender and associations through surface forms

The view that surface forms of grammatical gender forms heighten gender associations is also relevant to discussions that seek to alleviate these robust male biases. As will be discussed in the present thesis, grammatical forms can be manipulated to linguistically dissociate specific role nouns from a specific gender. Consideration of these issues is relevant as they may impact more tangible social issues.

In recent years, feminist linguists have actively promoted gender-fair language that makes both sexes linguistically visible (Duden, 2009) in an effort to overcome the male-dominating effects seen in language use. Possible means of achieving this goal involve explicitly emphasizing the presence of women through pair forms (e.g., German: *Lehrer*_{Masculine} *et Lehrerinnen*_{Feminine} [male and female teachers]) and capital I (only in German, e.g., *LeserInnen* [readers]), or neutralizing the context using non-differentiating forms (e.g., German: *die Angestellten* [the employees]).

Comparison of masculine forms to these alternative forms has revealed that they both result in different representations. In a series of experiments examining the effects of the masculine forms in French, Brauer and Landry (2008) asked participants to list potential politicians that would make for a good prime minister, historical or current heroes, favorite singers and actors. The authors found that when asked these questions using more inclusive forms (i.e., pair forms), participants' answers reflected a greater inclusion of women than when asked using the masculine forms. Importantly, the impact of these gender-fair forms have affirmed that they not only come into effect on a representational level, but also augment the level of perceived self-efficacy for certain occupations (Chatard et al., 2005) and vocational aspirations (Vervecken, Hannover, & Wolter, 2013) among children.

Research has also revealed other gender biases inherent in grammatical languages that may potentially counter the male-biases elicited by masculine forms. Although evidence is still unclear, these biases have been found to be triggered in a similar way to masculine forms; through surface forms of specific grammatical features. Contrary to the consistent reports of male-biasing effects of masculine forms, Rothermund (1998) identified an opposing but modest female bias embedded in the German plural determiner *sie* that reduced these male-dominant biases. In his study, participants were presented with short texts including person references in the singular (e.g., *der Student* [the student]) or plural masculine form (e.g., *die Studenten* [the students]) interpretable as generic. Recognition tasks presented afterwards

suggested that rejection times to test phrases for scenarios with masculine plural references decreased rejection times for sentences with female associations. The author claimed that these effects were indicative of a female bias, arguing the effect was instigated by the definite plural determiner in German (i.e., *die*), which shares the same surface form as the definite singular feminine determiner (i.e., *die*).

Together, these findings provide converging evidence to suggest that the mechanisms underlying the comprehension of gender information among grammatical gender language readers are complex. While readers may also rely on stereotypical information (as natural gender language readers do), grammatical gender information influences gender representation based on formal grammatical rules as well as through association of its surface forms that may outweigh stereotypical information. These effects of gender representation were found to be substantial, with far-reaching implications that touch on larger social cognitive issues.

These studies highlight that typologically different languages (i.e., natural gender languages vs. grammatical gender languages) convey gender information in various ways that result in diverging representation outcomes. This begs the question as to how readers of more than one language (e.g., bilinguals) would thenceforth mentally represent gender, and more importantly, if the performance tendency for gender representation may be altered as a function of language switches. Intuitively speaking, bilinguals of typologically different languages should change representation according to the language switches they undergo during reading – natural gender languages such as English evoking more stereotypical representations and grammatical gender languages such as French and German evoking more male biased representations. If this were the case, bilinguals would serve to provide insight into understanding how languages govern thought processes.

In the following, the discussion will shed light on the issue of the linguistic relativity hypothesis (Sapir, 1985; Whorf, 1956) and thinking-for-speaking hypothesis (Slobin, 1996a, 1996b, 2003) that postulates these possible effects of language on cognition, followed by a detailed account of bilingual research associated with the influence of language on cognitive processes.

2.5 The linguistic relativity hypothesis and the thinking-for-speaking hypothesis

The idea that language is responsible for our non-linguistic cognitive processes is known as the *linguistic relativity hypothesis*. Initially proposed by Sapir and his student Whorf, the underlying idea suggests that, as languages differ in the linguistic properties and expressions they impose on the speaker, they eventually enforce speakers of the language to frame their experiences and thoughts based on the linguistic properties emphasized therein (Sapir, 1985; Whorf, 1956). While the strong version of the view that language plays a deterministic role in cognitive processes is no longer supported, a growing body of research still supports the notion that language influences thought processes or guides cognitive processes and thus seeks to define the very specific conditions in which these effects can occur (e.g., Boroditsky, 2001; Brysbaert, Fias, & Noël, 1998; Vigliocco, Vinson, Paganelli, & Dworzynski, 2005; Winawer et al., 2007).

For instance, most studies in favor of linguistic relativity work under the assumption that different labeling practices, such as in color naming, lead to perceptive differences, in this case, on perception of the color spectrum (the same principle also extends to issues such as object categorization: Flaherty, 2001; Kurinski & Sera, 2010; mathematics: Pica, Lemer, Izard, & Dehaene, 2004; and person cognition: Curt Hoffman, Lau, & Johnson, 1956).

Notwithstanding the growing number of research in favor of this general idea, the difficulty in providing compelling empirical evidence for linguistic relativity still stands. Opponents of the hypothesis suggest there are relatively difficult confounds empirical studies face. That is, general cognitive mechanisms are highly complex, involving multiple cognitive processes co-occurring and influencing each other. It is hence reasonable to assume that they may co-vary with other variables (e.g., cultural variables: Swoyer, 2011).

Alternatively, many opponents argue that empirical evidence in support of linguistic relativity merely present “language-on-language” effects. The argument states that patterns of cognitive performance that reflect certain language-specific characteristics are simply a byproduct of language processing that is needed for problem solving mechanisms such as memory retention (Gleitman & Papafragou, 2013). In other words, language is inevitably a processing medium that allows people to engage in higher cognitive processes. It would thus be reasonable to assume that certain cognitive processes may reflect certain language-specific characteristics given an increased advantage of attention allocation that language processing can offer.

Franklin, Clifford, Williamson and Davies (2005) evaluated linguistic relativity focusing on color categorization. They argued that linguistic differences observed among adult speakers of different languages were arguably not perceptual, but may simply have resulted from verbal labeling that heightened an existing category. Their claims were based on developmental findings in which the perception of color categorization did not differ among English and Himba children despite different linguistic color labeling between the two languages. The lack of differences between color categorization showed that at least at pre-linguistic stages, color categorization can be considered as being universal. Their results also correspond to Lucy and Gaskins's (2003) study on entity categorization which demonstrated that performance in a sorting task diverged for English and Yucatec children. Results reflected language specific characteristics after 9 years of age, although differences were not observed before these linguistic differences were acquired (before 9 years). These studies argue that acquiring a set of linguistic rules allows perceptual mechanisms to be structured, although basic cognitive mechanisms may be universal regardless of the languages people acquire.

As these discussions establish, these language-on-language effects should be well distinguished from the influences of language on cognition *per se* which do not necessarily require processing through linguistic means (Papafragou, Hulbert, & Trueswell, 2008). Nonetheless, such so-called language-on-language effects should not be overlooked, given that most basic cognitive mechanisms undoubtedly require linguistic reflection to sustain cognitive efficiency, even if it is an unconscious stream of thinking.

In fact, further to this debate on language-on-language effects, these research findings and their implications have recently been clarified in the context of Levelt's (1989) speech production model which illustrates how language-specific characteristics become encoded within a communicated message. Based on this model, speech production is comprised of three main components. The first component, *conceptualization* suggests that the message to be transmitted needs to be determined by information stored in the speakers' memory. The second component *formulation*, involves devising the message into a specific linguistic expression and finally, the third component *articulation*, involves planning the phonetic delivery and the actual execution of the message. It is the second component *formulation* that would then assume that language-specific characteristics would be imposed on the message itself. Although Levelt (1989), along with other linguists such as Pinker (1989), characterized such linguistic formulation effects as being too unsubstantial to impact conceptualization after

speech production, his model has been adapted by Slobin (1996a, 1996b, 2003) as a basis on which to develop the *thinking-for-speaking* hypothesis.

The thinking-for-speaking hypothesis essentially assumes that typological characteristics of different languages bring speakers' attention to specific linguistic features. Formulating a verbal message requires speakers to select precise linguistic characteristics given that an event or message can be expressed in multiple, different ways. The specific way in which a speaker chooses to encode a message or an event hence constrains the message within the linguistic restrictions in which it is expressed. In other words, different languages, with their individual grammatical and lexical properties, should activate specific characteristics when speakers prepare themselves to speak. A message formulated in English would consequently differ from the same message formulated in French or German, given that there are language-bound characteristics encoded within the same message. The activation of such language specific characteristics can essentially bias the construction of mental representations. In this sense, certain linguistic characteristics of a language can maintain prominent effects during the encoding processes, which in turn influence attention and memory processes of the speaker.

Counter to the ongoing debate as to whether language may impact cognition, Slobin (1996a, 1996b, 2003) stressed the significance of differentiating between the idea that cognition is malleable and open to restructuring by language influences, and that of language influencing on-line construction of mental models. He argued, "We are not concerned with real world cognition here, but rather with the ongoing construction of mental representations." (Slobin, 2003, p. 260). For instance, in order to speak Spanish, English speakers need to learn the differentiation of honorifics in *tú* (you [singular and familiar]) and *usted* (you [singular with honorifics or plural with honorifics or just plural]) which are not expressed in the English equivalent *you* (example taken from Slobin, 2003). Speaking Spanish consequently compels the speaker to make note of the hierarchical differences between the speaker and their interlocutor.

In the context of this thesis, this idea relates to how gender information becomes (un)consciously assessed (or ignored for that matter). Whereas English speakers may refer to a person without having to specifically denote or attend to their sex (e.g., *the nurse*), for French speakers, the specification of gender to a human reference is obligatory through grammatical gender (e.g., *l'infirmière*_{Feminine} [female nurse]). Acquisition of French for English speakers requires them to learn means to consistently monitor the sex of persons being referred to. Efforts to make sense of incoming linguistic information would thus draw

speakers' attention to gender-specific information. This line of reasoning was central to the present thesis, which aimed to substantiate this principal argument stressed by Slobin (1996a, 1996b, 2003) that the significance of the impact of language on language process lies in the idea of *language in use*, and that the mental activities associated with formulating an utterance during speaking, reading or listening deserves further exploration.

Much support for the thinking-for-speaking hypothesis has been gained through investigation of how a specific experience or an event is encoded in language. A common example is found in motion event expressions that assume typological differences in how motion is lexicalized in verbs between satellite and verb-framed languages. Satellite languages such as English are known to encode the path of motion in associated particles of the verb (e.g., The dog went *into* the house.) and manner of motion directly in their verbs (e.g., The dog *ran* into the house). On the other hand, verb-framed languages such as French rely more on the verb itself to code path of motion (e.g., Le chien est *entré* dans la maison. [The dog *entered* the house.]⁷), although manner of motion is not coded directly on the verb (e.g., Le chien est entré dans la maison *en courant*. [The dog entered the house *by running*.]; example taken from Slobin, 2003).

In their formative study, Berman and Slobin (1994) investigated how linguistic production of children and adults of verb-framed and satellite languages would differ. They asked participants from various languages in verb-framed (e.g., Spanish) and satellite languages (e.g., English) to narrate stories from a picture book illustrating different events. They found that overall, English speakers were more likely to use verbs describing manner of motion than the speakers of verb-framed languages. They concluded that production differences resulted from speakers' choices to code certain information, which was essentially guided by their language.

Similarly, Gennari, Sloman, Malt and Fitch (2002) examined how differences between Spanish and English speakers as regards lexicalizing motion events influenced speakers' non-linguistic tasks. They presented English and Spanish participants with movie clips of motion events (e.g., an agent walking into a room) and asked them to perform a recognition memory task followed by a similarity task (i.e., making judgments as to whether the movie clips showing the movements were the same or not). Conditions varied across participants during the event presentation in which they were asked to either simply watch the events or repeat

⁷Although this direct translation is grammatically correct in English, Slobin (2003) suggests this phrasing would not be the "habitual" means to encode the sentence which the thinking-for-speaking hypothesis considers as a key issue.

nonsense syllables while watching the clips. They found no effect of recognition memory in both linguistic and non-linguistic encoding conditions, although the similarity task conducted after linguistic encoding showed performance consistent with participants' linguistic pattern. These results allowed for a differentiation between linguistic and non-linguistic tasks, leading the authors to propose the language-as-strategy view. This idea suggested that language is used to mediate judgments for specific tasks on-line, which consequently mirrors language-specific patterns.

While much work has substantiated accounts of the effects of specific language-bound characteristics on linguistic processing, another avenue of research investigating linguistic effects on cognitive processes has been to investigate *bilinguals*, a population sample which the present thesis focuses on. Indeed, Kousta, Vinson, and Vigliocco (2008) argue that a comparison between bilinguals and monolinguals allows for an investigation of the magnitude of language effects on cognition to be made. Ideally, bilinguals could provide a means of tackling whether modes of thought differ or *shift* according to the different languages people apply to encode specific concepts or ideas.

2.6 Testing the influences of language among bilinguals

In general, bilinguals are considered to have competencies in two languages: a native, dominant language (L1) that is acquired from birth onwards, and a second language (L2) that is acquired subsequently or simultaneously. In L2 research examining thinking-for-speaking effects, bilinguals' abilities in the two languages are manipulated by examining the different performances in experimental tasks in each of their languages. It is assumed that generating inferences for example, may result in different language bound performance biases.

One of the most influential models of bilingual lexical processing known as the Revised Hierarchical Model as shown in Figure 2.1, (Kroll & Stewart, 1994; Kroll & Tokowicz, 2005) provides a framework to approach the underlying mechanisms of how L1 and L2 may interact during language comprehension. According to this model, both the L1 and L2 lexicons are connected respectively to the conceptual store. The strengths of the links however are not equal, with L1 having a stronger connection than L2 to the concept. Furthermore, there is a link between L1 and L2, where L2 lexical items are more strongly connected to their L1 equivalents than L1 lexical items are to their L2 equivalents. This implies that during early L2 acquisition, learners will rely more on their L1 to access the meaning of newly learnt L2 lexical items.

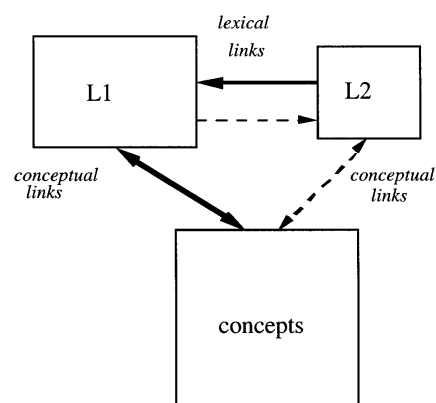


Figure 2.1 The Revised Hierarchical Model (Kroll & Stewart, 1994; Kroll & Tokowicz, 2005)

In cases where learners need to learn new L2 rules that are non-existent in L1, they will need to adapt or rely heavily on L1 regularities to form new L2 rules (MacWhinney, 1997). Naturally, the regularities of L1 will emerge as linguistic “accents” (e.g., semantic, syntactic, etc., MacWhinney, 1997) even when processing in one’s L2. In contrast, as a better understanding in the L2 is developed, the reliance on L1 will consequently lessen, showing greater autonomy from L1. L2 regularities should thus become more evident during the comprehension process, showing a shift from L1 to L2 reliance for comprehension.

One line of evidence that supports the notion of bilinguals’ processing shift comes from research examining change in categorization tendency according to the language applied to encode information. For instance, in studies investigating bilinguals’ color categorization, Athanasopoulos (2009) examined how the color blue was cognitively represented in Greek-English bilinguals. While the Greek language differentiates blue into darker and lighter shades, a similar distinction is not characteristically made in English. They found evidence suggesting that bilinguals’ preference for color differentiation diverged from their L1 towards that of their L2 tendency. Furthermore, it was revealed that this effect in performance bias shifted in conjunction with the participants’ level of bilingualism (determined by their acculturation level and length of stay in the L2).

In a similar vein, Athanasopoulos and Kasai (2008) investigated how grammatical effects of number influenced categorization of novel objects in Japanese-English bilinguals, as the two languages diverge in their linguistic tendency to categorize objects. While in Japanese, speakers commonly categorize objects based on material, for English speakers, there is a preference for object categorization to be based on shape. Japanese-English bilinguals presented categorization preferences of their L2 with L2 proficiency (controlled for socio-cultural factors, length of stay in the L2 context and general L2 proficiency tested by

the Oxford Quick Placement Test) being a vital predictor of the extent these performance biases presented. Those with high L2 levels performed like native speakers of participants' L2 English, categorizing objects based on shape. In contrast, intermediate learners of L2 English relied more on their L1 Japanese, and were more likely to categorize the objects based on material.

More pertinent to the issues of this thesis, Danziger and Ward (2010) showed that shifts in bilinguals' languages could even alter activation of stereotypical biases associated with speaking a specific language. In their study, they predicted that Arabic-Hebrew bilinguals' implicit predispositions associated to using the Arabic language would lead to the activation of more positive associations to Arabs than to Jews than when using Hebrew, given that Arab Israelis are a minority as opposed to Jews within the community. Language was essentially expected to activate associated traits. The authors presented their bilingual participants with an IAT (implicit association test) that required them to immediately categorize words with a button press. Positive (e.g., *clean*, *good*) and negative (e.g., *mean*, *weak*) words were paired with Arab (e.g., *Ahmed*) or Jewish (e.g., *Yair*) names for each button. Participants' performance showed that responses were significantly faster when Arab names were paired with positive associations and Jewish names with negative associations, although the positive effect for Arab names was significantly greater when the task was conducted in Arabic. Using the same line of reasoning and the IAT, convergent results were reported by Ogunnaike, Dunham, and Banaji, (2010) employing Arabic-French and Spanish-English bilinguals.

As reflected in these studies, although the influences of language on cognition *per se* are still subject to debate, in general, there appears to be a common agreement that language has a substantial impact on cognitive processes that incorporate both linguistic and nonlinguistic processing. These studies suggest that the activation of certain languages may emphasize particular language-specific characteristics for the perceiver, which in turn will highlight concepts that may influence subsequent processing. Moreover, processing in a specific language may also serve to activate relevant information that is linguistically associated (i.e., stereotypical biases).

Interestingly, for bilinguals who maintain two languages as part of their language system, this means that the emphasis towards specific regularities will differ according to each language, and their diverging tendencies may pose shifts in mental representations to occur. However, the extent to which these language performance tendencies may influence further processes appears to be dependent on factors associated to the level in which language

is rooted in the cognitive system (e.g., Athanasopoulos & Kasai, 2008; Athanasopoulos, 2009), in other words, bilinguals' *proficiency*.

2.7 Task performance and L2 proficiency

While most laypersons may consider bilinguals as those with a native-like mastery of two languages, the definition in the present thesis applies a more general definition adopted in most bilingual research: *a person who regularly uses two languages*. Defining which L2 competencies constitute a bilingual is difficult, as the abstract notion of language competency spans a continuum. However, the notion of L2 competency, or in more general terms proficiency, has been inseparable from studies investigating linguistic processing in bilinguals, given that proficiency has direct effects on the performance of experimental tasks. Indeed, as will be addressed in this thesis, studies affirm that the magnitude of linguistic effects over cognitive processes may very well differ in function to bilinguals' proficiency levels.

For example, Phillips and Boroditsky (Experiment 2, 2003) examined how L2 proficiency in Spanish-German bilinguals varied the extent to which images of objects and people were perceived as being similar. They predicted that if grammatical gender of the presented objects coincided with the presented image of the person's gender, similarity ratings between the two would be high. The authors did indeed find evidence of such an effect, with participants' L2 German proficiency proving as a reliable indicator as to the extent that bilinguals' responses conformed to the grammatical gender system. While these effects were arguably attributed to participants' L2 proficiency, evaluation of L2 proficiency in their study was based solely on a self-assessment ranging from 1 (*not fluent*) to 5 (*very fluent*). A questionable criterion, which neither sufficiently nor objectively captured the bilinguals' proficiency levels.

In contrast, Athanasopoulos and colleagues have provided a more comprehensive picture of how L2 proficiency contributes to the impact of language on cognitive effects by using a more diverse evaluation method. In their aforementioned studies examining color and object categorization, Athanasopoulos and colleagues (e.g., Athanasopoulos, Damjanovic, Krajciova, & Sasaki, 2011; Athanasopoulos & Kasai, 2008; Athanasopoulos, 2009) have consistently demonstrated that L2 performance matched specific linguistic regularities, with L2 proficiency being a good indicator of the magnitude of which these effects emerged. In their studies, L2 proficiency was gauged through different objective criteria, ranging from

standardized tests (e.g., Oxford Quick Placement Test, Nation Vocabulary Test) and participants' individual language background such as the length of stay in an L2 speaking environment or frequency of L2 usage.

While these studies depict a good picture of the effects of L2 proficiency, they are not informative as regards the impact of proficiency on reading comprehension which this thesis focuses on. Importantly however, L2 proficiency along with reading abilities have been identified as two essential variables that contribute to good and poor reading (e.g., Alderson, 1984; Carrell, 1991; Clarke, 1979). This idea has been framed within the short-circuit hypothesis (nowadays commonly known as the linguistic threshold hypothesis) that identifies L2 readers' dependency on bottom-up processes (Clarke, 1979). It assumes that given the lack of L2 knowledge, much of L2 readers' resources will be allocated to low-level processing that involves mostly decoding surface forms (i.e., lexical and syntactic forms). Consequently this allocation of resources leads to restricted top-down processes that may limit the depth of understanding of the text as a whole (Zwaan & Brown, 1996). Clearly, these basic bottom-up processes are automatic (or at least facilitated) during L1 comprehension which amounts to different reading comprehension behaviors between the two language groups. Clarke (1979) proposed that L2 readers need to attain a minimum level of L2 proficiency before comprehension processes can be effectively carried out (Clarke, 1979). In contrast, the latter position (that views general reading abilities as essential) is described by the linguistic interdependence theory (Alderson, 1984), and assumes that there is a transfer of L1 reading abilities to the L2. Such an account assumes that new reading strategies do not necessarily need to be acquired in the L2 and that good comprehenders in their L1 are essentially good comprehenders in their L2.

While these theories appear to be rather exclusive, Bernhardt & Kamil (1995) posit that there lies a considerable amount of consistency between the two theories. For example in Zwaan and Brown's (1996) study, English learners of French read texts in their L1 English and L2 French while performing a think-aloud task. Results indicated differences between L1 and L2 reading, with processing in the L1 resulting in construction of stronger situation models (indicative of deeper levels of processing). On the other hand, L2 processing revealed more reliance on low-level processing (i.e., vocabulary, semantic access and syntactic strategies) resulting in limited comprehension. These results suggest that L2 proficiency is key in realizing basic surface level decoding. If this process consumes most resources, higher level processing would be limited, ultimately leading to poor comprehension.

In a more recent study, Horiba (2000) compared L1 and L2 reading for different texts (stories and essays) in conditions where participants were instructed to read freely or for coherence. Results for the L2 readers revealed that, although they showed different types of processing tendencies according to text type (e.g., more elaborative inferences on text structure were generated for essays than stories), these effects were based on a general L2 reading inefficiency. In other words, L2 readers allocated a greater level of processing capacity to process low-level information that resulted in a lack of processing resources to attend to higher levels of processing (e.g., information integration), confirming L2 reading problems relating to lack of language proficiency.

Together, these studies demonstrate that the extent to which language-specific processing biases take effect are heavily dependent on participants' language competence in the L2, as it appears to be a good predictor in the allocation of bottom-up and top-down processing during L2 reading and more importantly, shifts in mental representations. Surface levels cues in particular may inflict a greater magnitude in the L2 than would be in the L1 as L2 readers would attend to decoding surface level information.

With regard to the studies reported in this thesis, gauging participants' L2 proficiency was important in understanding the underlying processes of the effects that were being investigated. To do this, a type of cloze test known as the C-test was employed for all experiments in English, French and German. In the following, a review of the evaluation methods for L2 proficiency is provided with a detailed description on the use of C-tests.

2.8 Evaluation of L2 proficiency

A recent review by Hulstijn (2012) has highlighted the critical debates around evaluation of language proficiency in L2 research, revealing that assessment of participants' proficiency to date has been insufficient as regards grasping the real nature of language competence. As seen in the effects of previous studies described earlier, participants' L2 proficiency has a direct impact on the effects that are being investigated and should be rigorously tested. Nonetheless, attaining a unique participant sample in a research context in terms of L2 proficiency has been difficult, given the significant variability in what constitutes L2 competency.

To name a few of these contributing factors, the most obvious factors are age of acquisition, length of study, context of L2 study (e.g., natural vs. classroom setting) and frequency of usage. Nonetheless, these factors may also differ according to the social

dynamic, for instance, learning an L2 in the United States where the social mixture of cultures and languages is relatively dynamic, would be much different to learning an L2 in a country like Japan where the population is relatively homogenous. Such intricate dynamics suggest that it is difficult to fully account for these variables in research.

Past studies employing bilinguals have administered various assessment tests for *objective* L2 evaluation. For instance, participants have been screened based on standardized test scores from the Oxford Placement Test (Clahsen, Balkhair, Schutter, & Cunnings, 2013), Nation's Vocabulary Test (Athanasopoulos, 2009; Athanasopoulos et al., 2011), Oxford Quick Placement Test (Athanasopoulos & Kasai, 2008), Peabody Picture Vocabulary Test (Prior & MacWhinney, 2010) and the verbal component of the SAT (Prior & MacWhinney, 2010). Many have also coupled these objective tests with participants' self-assessment data (Lehtonen & Laine, 2003). In particular, there has been a growing interest in the use of C-tests, in other words, a renewed format of a type of cloze test.

The execution of the C-test requires participants to fill in blanked out gaps in a given time frame. Every second half of every other word of a text is blanked out except the initial and final sentences of the text. Knowledge associated with reconstructing the blanks is considered to show participants' ability to anticipate and construct language (Grotjahn, Klein-Braley, & Raatz, 2002), with the number of correct restorations used as the indicator of proficiency.

Initially, while earlier research used cloze tests (i.e., fill-in-the-blank type of testing), criticism regarding its procedure was raised. In terms of the testing format, the major critique was that the texts were too long (as they needed to provide enough blanks to meet the deletion requirements). Furthermore these arbitrary rules of deletion principles made it difficult to maintain a level of reliability. These concerns were subsequently addressed, leading to the creation of the C-test (initially devised by Raatz and Klein-Braley, 1982), which is now seeing a gradual increase in implementation in L2 research.

More recent discussion on the implementation of C-tests has seen a more focused trend on examining what C-tests actually measure. In particular, studies have reported correlations between the C-test and the principal language competencies (e.g., vocabulary, oral abilities, speaking, listening and writing), while sophisticated means to evaluate their content using Rasch modeling and confirmatory factor analyses (e.g., Eckes & Grotjahn, 2006) have further confirmed their reliability to measure L2 proficiency.

2.9 The bilingual influence on gender representation

Thus, if, as previous studies show, there are typological differences between how readers of natural gender and grammatical gender languages construct the mental representation of gender, there is a great discussion to be had on a social and cognitive level in terms of gender biases that are evoked by languages.

There is a need to reveal the basic mechanisms underlying how certain gender biases are evoked and represented, in as much as it is tantamount to revealing how languages in general influence our cognitive processes. In pursuing such a goal, the following chapters (Chapters 3 - 6) of this thesis examine bilinguals' gender representation processes as they provide an insight into how gender biases may be shifted and altered on a cognitive level. In particular, given that the debate on L2 proficiency and language processing suggests that the influence of bilinguals' L2 proficiency may serve as a key element in such a discussion, L2 proficiency is further dealt with throughout the studies presented onward. We now turn to the experimental studies of the thesis.

Chapter 3 : Study I

Gender inferences: Grammatical features and their impact on the representation of gender in bilinguals

The study presented in this chapter was published as:

Sato, S., Gygax, P. M., & Gabriel, U. (2013). Gender inferences: Grammatical features and their impact on the representation of gender in bilinguals. *Bilingualism: Language and Cognition*, 16, 792–807. doi:10.1017/S1366728912000739

3.1 Abstract

In this study, gender representation processes of French-English and English-French bilinguals were investigated based on the assumptions of Gygax et al. (2008) who differentiated the representation tendencies between native speakers of the two languages. Representational shifts were expected to surface with language shifts given that the linguistic regularities of each language was expected to impose readers to attend to specific information – French on male frames of reference and English on gender stereotypical frames of reference.

We therefore investigated the effects of grammatical and stereotypical gender information on the comprehension of human referent role nouns among bilinguals of a grammatical (French) and a natural gender language (English). In a sentence evaluation paradigm, participants judged the acceptability of a gender-specific sentence referring to either a group of *women* or *men* following a sentence containing the plural form of a role noun female (e.g., *social workers*), male (e.g., *surgeons*) or neutral (e.g., *musicians*) in stereotypicality. L1 French and L1 English bilinguals were tested both in French and English. The results showed that bilinguals construct mental representations of gender associated with the language of the task they are engaged in, shifting representations as they switch languages. Specifically, in French, representations were male-dominant (i.e., induced by the masculine form), whereas in English, they were stereotype-based. Furthermore, the results showed that the extent to which representations shifted was modulated by participants'

proficiency in their L2, with highly proficient L2 participants resembling native speakers of the L2 and less proficient L2 participants being influenced more by their native language.

3.2 Introduction

Text comprehension involves generating inferences from textual information and one's world knowledge in order to create a coherent mental model that is representative of the depicted textual information (Garnham & Oakhill, 1996; Johnson-Laird, 1983; Zwaan & Radvansky, 1998) and also in line with readers' world knowledge.⁸ Proponents of constructionist accounts (e.g., Graesser, Singer & Trabasso, 1994) stipulate that readers undergo a process in which information such as the spatial situation, the person's emotions, their physical attributes, or the story goals are inferred in order to fill informational gaps. Of such inferences, those associated with the protagonists' gender have received some, yet quite sparse attention in the past 15 years.

Constructionists' view would suggest that readers rely on gender information stretching from gender stereotypes as part of world knowledge and experience, as well as language features such as grammatical gender to yield gender-associated inferences, and that these inference processes are *automatically* (Banaji & Hardin, 1996; Cacciari & Padovani, 2007; Irmen, 2007; Irmen & Roßberg, 2004; Oakhill, Garnham, & Reynolds, 2005) and *elaboratively* activated (Garnham, 2001; Garnham, Oakhill, & Reynolds, 2002).

In natural gender languages such as English, where gender-associated information is conceptually and semantically embedded and is not overtly marked on a grammatical level – except some obsolete terms such as *policeman* or *postman* – readers may need to resolve attributes like the gender of an encountered referent based on world knowledge. In such a case, readers may rely on stereotypical gender attributes to infer whether the protagonist is a man or a woman. For example, readers may infer that a protagonist is a woman given that she is depicted as being “sensitive” and “sympathetic”, which are considered common stereotypical characteristics of a woman (Hosoda & Stone, 2000), or through a definitional referent that she is a “sister” or a “mother”.

The process of activating such gender-associated information is a relatively complex top-down one, and to investigate the effects and activation of such inference generation

⁸ Unless otherwise stated, the terms speakers, readers, comprehenders and participants are used interchangeably to indicate a group of people and/or participants of a certain native language group.

processes, studies have employed occupational role nouns associated with gender stereotypes (e.g., male stereotype: *firefighter*, female stereotype: *secretary*, neutral stereotype: *artist*) in experiments using *match-mismatch* paradigms (Carreiras, Garnham, Oakhill & Cain, 1996; Garnham et al., 2002; Irmen, 2007; Kreiner, Sturt & Garrod, 2008; Oakhill et al., 2005; Reynolds, Garnham & Oakhill, 2006). Occupational role nouns normally do not denote a specific gender, but in many cases have a strong association to a specific gender depending on the likelihood of occurrence in the real world (i.e., nurses are more frequently women than men). Thus, unless stated otherwise in the context, readers need to rely on information from their world knowledge or discourse cues to infer the gender of the referent. The match-mismatch paradigms take advantage of the fact that when information that is inconsistent with these gender stereotypical beliefs is presented following a stereotypical occupational role noun, a mismatch effect in the dependent variable (e.g., longer reading times or eye-fixation times) would surface, indicating difficulty in integrating the information into readers' mental representations.

In a series of experiments conducted in English, Oakhill et al. (2005) found that when being asked to judge whether a gender stereotypical occupational role noun (hereafter referred to as *role noun*) and a kinship noun with a semantic gender (e.g., male: *uncle*, female: *sister*) could represent the same person, participants exhibited a mismatch effect reflected in increased judgment times when the gender stereotypicality of the role noun and the semantic gender of the kinship noun did not match. They interpreted these results as reflecting the immediacy of participants generating gender stereotypical inferences, thus initiating increased judgment times to overcome comprehension difficulty and to update their representations. These immediate inferences were also difficult to suppress, even when participants were explicitly instructed to do so, suggesting the robustness of the nature of generating such inferences.

Interestingly, these mismatch effects were not only observed in lexical-level tasks, but also during sentence comprehension (Carreiras et al., 1996; Duffy & Keir, 2004; Garnham et al., 2002; Kreiner et al., 2008). Duffy and Keir (2004), for example, reported increased eye-fixation times on reflexive pronouns (e.g., *himself* / *herself*) that followed a mentioning of role nouns (e.g., *electrician*) presented in an earlier sentence that mismatched the pronoun in gender stereotypicality. Similar findings were also reported by Carreiras et al. (1996), corroborating the notion of an automatic activation of gender stereotypical information and

the difficulty to map mismatching gender information onto readers' mental representations constructed during text comprehension.

Overall, these studies addressed the fact that readers automatically activate gender information of a referent if stereotypical information is readily available in the text. These experimental tasks have been conducted primarily in English, a natural gender language. Therefore, not surprisingly, stereotype-related information as an influence on gender representation and inferences was of primary concern. Fundamentally, there should be little reason for these effects not to be seen in other languages, but recent evidence suggests that gender-associated inferences are not only triggered by available stereotypical information but also by grammatical features (i.e., grammatical gender) found in grammatical gender languages (Cacciari, Corradini, Padovani, & Carreiras, 2011; Cacciari & Padovani, 2007; Carreiras et al., 1996; Gygax & Gabriel, 2008; Gygax, Gabriel, Sarasin, Oakhill, & Garnham, 2008; Irmen, 2007; Irmen & Roßberg, 2004; Stahlberg, Sczesny, & Braun, 2001).

In grammatical gender languages such as French, German or Spanish, both animate and inanimate nouns are morphologically marked for gender (e.g., masculine, feminine, neuter). Inanimate nouns are categorized arbitrarily according to each language, such as *la chaise*_{French} (Feminine: the chair) or *der Bleistift*_{German} (Masculine: the pen], but in cases of personal nouns, the grammatical and biological gender of the person typically correspond, as in *une étudiante*_{French} (Feminine: a female student) and *un étudiant*_{French} (Masculine: a male student). This rule of gender categorization in grammatical gender languages is fundamental, as verbs and adjectives are inflected for gender agreement in relation to these nouns. Researchers generally agree that the acquisition of grammatical gender in the first language (L1) occurs effortlessly (Karmiloff-Smith, 1979; Lyster, 2007), and native speakers have shown to be highly sensitive to grammatical cues of gender in studies investigating syntactic gender violations (Barber & Carreiras, 2005), interaction of semantic and grammatical gender information (Wicha, Orozco-Figueroa, Reyes, Hernandez, Gavaldón de Barreto & Bates, 2005) and within conceptual processing investigating perceived masculinity / femininity of objects and entities (Bassetti 2007, 2011; Sera, Elieff, Forbes, Burch, Rodríguez & Dubois, 2002), suggesting that grammatical gender categorization plays a crucial role in language processing. In particular, in a series of experiments employing a sentence-picture semantic judgment paradigm, Wicha et al. (2005) found an interaction for grammatical gender and semantic congruity. In their first experiment, for example, participants were asked to name a presented picture, which replaced a critical noun of an auditory presented sentence. The

critical nouns were always preceded by a congruent or incongruent gender-marked article, and were semantically congruent or incongruent with the context. Naming times showed that even subtle grammatical gender information (in their case, gender-marked articles) along with semantic information (i.e., congruency of contextual information and the presented picture) contributed to sentence meaning and influenced speakers' subsequent language production. Given such effects of grammatical gender features on language comprehension and the relative automaticity of gender stereotyping, numerous studies on the construction of gender representation have argued that these two sources of information intricately interact when both are readily available to the comprehender.

As mentioned earlier, although grammatical gender for human references specifically marks for their biological gender, the masculine form can also be used in conditions where the sex of the person(s) is unknown, irrelevant or is mixed, and is intended in a generic sense, producing ambiguity as to a male-exclusive or a generic interpretation. Though still limited, psychological studies have provided compelling evidence in grammatical gender languages showing that the usage of the masculine plural form intended as a generic interpretation evokes less female representations than its female counterparts as well as other gender neutral forms (e.g., pair forms: *étudiants / étudiantes*) by favoring stronger associations with the male gender (e.g., German: Braun, Sczesny, & Stahlberg, 2005; Irmen & Köhncke, 1996; Stahlberg et al., 2001; French: Brauer & Landry, 2008; Gygax et al., 2008; to some extent Norwegian: Gabriel & Gygax, 2008).⁹

Consequently, the use of the masculine form as generic may have serious implications when put into a real- world context. For example, Vervecken and Hannover (2012) found that when presenting traditionally male-held jobs in pair forms in German (i.e., presenting both masculine and feminine forms together), children of 6–13 years of age were more likely to access female representations. Most importantly, the children considered women as being potentially more successful in these jobs when presented in both the masculine and feminine forms (Study 1 and 2). Similarly, Chatard, Guimont and Martinot (2005) investigated how the usage of masculine, feminine and epicene forms influences the degree of self-efficacy toward

⁹ Studies are also present in English where presenting job applications (Bem & Bem, 1973; Stericker, 1981) in the generic form influenced children's aspirations and expectations (Liben, Bigler, & Krogh, 2002), increasing both the visibility and perceptions of women. Most importantly, these studies denote the difficulty of inferring intended generic interpretations even in a natural language where the gender is not marked on a grammatical level.

occupational role nouns among French middle school pupils. The pupils were generally more confident to uptake jobs that were stereotypically congruent with their own sex. However, for occupations that were stereotypically incongruent, the usage of female-inclusive forms (i.e., feminine form: *Mathématicien(ne)* [mathematician], epicene: *Mathématicien / Mathématicienne*) augmented the self-efficacy scores significantly more (especially for girls) than when the masculine form (i.e., *Mathématicien*) was used. These findings suggest that the linguistic means to present gender-associated information considerably influences the attitudes of the comprehender. Thus, evoking less female representations may disadvantage women in contexts where women are under-represented in terms of a male-dominant job market and may decrease people's vocational aspirations and occupational self-efficacy (Vervecken & Hannover, 2012).

Many studies associated with the generic interpretation of the masculine form have been derived from offline studies in which participants are given the opportunity to reflect on their beliefs and perceptions, rather than from online task studies where clearer interpretations of the type of information being immediately activated can be made. Gygax et al. (2008) adapted an online sentence judgment task to look at the interaction of grammatical gender and stereotypical information on the interpretation of the masculine form used as a generic in grammatical gender (i.e., French and German) and natural gender (i.e., English) language speakers. In all three languages, speakers were presented with sentences with a role noun, either female (e.g., *dressmakers*), male (e.g., *spies*), or neutral (e.g., *pedestrians*) in stereotype (selected from a cross-linguistic norming study by Gabriel et al., 2008). English participants were more likely to judge succeeding sentences as *sensible continuations* when they included gender continuations (e.g., *a group of men*, *a group of women*) that matched the gender stereotypicality of the role noun. In French and German, participants were more likely to accept sentences mentioning male continuations (i.e., *a group of men*) regardless of the gender stereotypicality of the role noun, indicating a male-dominant representation induced by the use of the masculine plural form (although intended as generic). The authors argued that stereotypical information was the active source for comprehension in English speakers due to the lack of a grammatical gender marking, whereas in French and German, the masculine form as *specific* (i.e., as opposed to *generic*) was the key source overriding available stereotypical information. Both natural and grammatical gender language speakers had the same stereotypical information available, yet grammatical information was a stronger determiner for gender representation among the grammatical gender language participants

than stereotypical information. In sum, the construction of gender representation on the mental model among grammatical gender language speakers may include the interaction of both top-down (stereotypical information from world knowledge) and constraining bottom-up (grammatical gender) influences, though the former seems to be overridden by the latter.

The findings of Gygax et al. (2008) document that speakers with certain language features (i.e., grammatical gender language or natural gender language) construct gender representations in different ways, suggesting different linguistic behaviors despite having the same stereotypical gender-associated information conveyed to them, and that it is essentially the linguistic source of which the gender-associated information derives from that fundamentally changes how these pieces of information are processed.

In light of these findings, bilinguals who speak two languages with different language features represent an attractive focus of research (Cook & Bassetti, 2010; Pavlenko, 2011). The complex relationship between constructed mental representations and gender-associated informational biases can be elucidated through a closer look onto processes inherent to language switch.

At this point, it is worth mentioning the differentiating notion of *conceptual* and *semantic* levels of representation. The former refers to the non-linguistic representation of an entity, whereas the latter refers to its linguistic components, including the speaker's knowledge of a word such as its definition or grammatical features / rules of a language. In activating concepts, certain semantic representations can impose different linguistic constraints, hence activating only distinct aspects of the concept (Paradis, 1997). Thus, depending on the language in which they are engaged in for comprehension, bilinguals may activate distinct conceptual components. Consequently, this could result in a shift in conceptual representation or in a cognitive restructuring. When second language (L2) competence and proficiency have not reached a sufficient level, the access to a concept in L2 is most likely controlled by an L2-to-L1 translation, resulting in language specificities that are more salient in L2 (e.g., linguistic features and characteristics that may or may not be existent in the L1) to be less influential.

Studies by Athanasopoulos and colleagues (Athanasopoulos, 2009; Athanasopoulos, Damjanovic, Krajciová, & Sasaki, 2011), for example, have revealed that bilinguals tend to shift their color naming categorization patterns to those of the native speakers of the target language or a pattern that falls to what they refer to as “in-between” the two patterns, and that these tendencies strongly influence the perception of distinguishing the actual colors.

Interestingly, and this is important in the present study, the degree to which this pattern was manifested was modulated by L2 proficiency which they defined as the frequency of the L2 use (exposure to the L2) (Athanasopoulos et al., 2011). Their findings mainly illustrate that advanced bilinguals can present cognitive flexibility in being able to behave in similar ways as native speakers of their L2. This flexibility mostly relies on linguistic, social and cognitive factors, relatively independent of the development of L1.

Findings in object categorization concur with this idea. Athanasopoulos and Kasai (2008) examined the notion that native English speakers have a disposition to categorize objects according to shape whereas native Japanese speakers show a preference for material, as the former language stresses the plural marking for count nouns in its grammar (e.g., *three apple- s*), whereas the latter does not (e.g., *san-ko no ringo* “three-piece of apple”). They found that unlike Japanese and English monolinguals, Japanese–English bilinguals manifested a categorization preference that differed from the monolingual tendency, and that changed according to their L2 proficiency. Advanced L2 speakers shifted their behavior toward L2 native patterns whereas intermediate L2 speakers remained close to their L1. The authors claimed that acquiring an L2 with grammatical concepts non-existent in the L1 could potentially reorganize cognitive structures in bilinguals and that the extent of the reorganization was modulated by L2 proficiency.

Together, the findings of the studies from color and object categorization speak to the idea that despite the fact that native cognitive patterns have already been established within a speaker, new cognitive patterns modulated by language proficiency could be acquired (Athanasopoulos et al., 2011). In this line of thinking, the present study primarily aims to question whether bilinguals infer gender differently when switching from L1 to L2. Secondly, it examines the influence and functionality of participants’ L2 proficiency as a modulating factor of this inference process.

Although an abundant amount of research has been conducted in the field of categorization (of different nature) among bilinguals, in which bilinguals were shown to be cognitively affected by specific linguistic features, to our knowledge, only a few studies on the way bilinguals represent gender have been conducted, and these were mainly focused on gender attribution (name attribution: Boroditsky, Schmidt, & Phillips, 2003; voice attribution: Flaherty, 2001) and gender agreement (Sabourin, Stowe, & de Haan, 2006; White, Valenzuela, Kozłowska-Macgregor, & Leung, 2004) in the target language. Still, some of the

studies on bilingualism are highly relevant to our present study, as they anchor the very hypotheses that we advance.

Kousta, Vinson and Vigliocco (2008), for example, asked monolingual Italian, English and bilingual Italian-English speakers to name pictures of animals that were presented at a fast rate. Under the premise that grammatical gender increases semantic similarity, the aim of the task was to elicit semantic substitution errors (answering *tiger* as opposed to *lion*). Results suggested that Italian-English bilinguals elicited more gender-preserving errors when the task was in Italian, than when the task was conducted in English. Interestingly, their behavior mirrored those of monolingual Italian speakers when conducting the task in Italian and the one of monolingual English speakers when conducting the task in English. The authors argued that the behavior of bilinguals could be predicted by the behavior of the native speakers of the target language, also known as intraspeaker relativity. Furthermore, given that conceptual representation is normally associated with L1, the authors claimed that their findings (i.e., bilinguals manifesting different behaviors in each language) supported the idea of language-specific effects on semantic representation (Pavlenko, 1999, 2011).

More pertinent to our research, Scheutz and Eberhard (2004) examined whether the German morphosyntactic ending *-er* – associated with the masculine gender in nouns, as in *Sprecher* (male speaker) – would automatically activate the similar *-er* denotation in participants' L2 English – associated with agentive nouns, as in *speaker*, but unrelated to the masculine gender. Their simulation data (and to some extent their eye-tracking data) confirmed that when reading a sentence with a stereotypical role noun (male, female or neutral) ending in an *-er* (e.g., *hunter*), followed by a reflexive pronoun (*herself*, *himself*) that referred back to the referent, German-English bilinguals elicited a male bias that was predicted by the morphosyntactic *-er* ending, whereas English monolinguals did not. They also attributed their weakened results to participants' age of L2 acquisition, which they associated with L2 proficiency, in line with what others have (Kim, Relkin, Lee, & Hirsch, 1997; Weber-Fox & Neville, 1996). This study is particularly important here for three reasons. First, it assessed the influence of L1 and L2 grammatical features on gender representation. Secondly, it showed that even rather minor morphosyntactic features could influence bilinguals' comprehension. Thirdly, it demonstrated that L1 and L2 processing may not be independent, which is in line with an interactive view of language processing (de Groot, Delmaar, & Lupker, 2000; Dijkstra, Grainger, & van Heuven, 1999; Dijkstra, Van Jaarsveld, & Brinke, 1998; Scheutz & Eberhard, 2004).

In all, though the studies on bilingualism presented so far have not come to an agreement on how L1 and L2 influence language processing, they provide compelling evidence to suggest that possessing more than one language can potentially and substantially affect how we comprehend certain types of linguistic information as well as the means in which we use it to build a mental representation of the world. Against this background, our research sought to extend Gygax et al.'s (2008) results by addressing possible shifts of representation within speakers of two languages. Overall, there were three hypotheses.

First, given that past results reported by Gygax et al. (2008) showed comprehension patterns to strongly differ between speakers of grammatical gender and speakers of natural gender languages, we expected to see a similar effect in terms of comprehension tendency in participants' L1 (i.e., male-dominant representation in French for French speakers and a representation in line with stereotypes in English for English speakers).

Second, we expected a shift of representation within each participant as they switch from one language to the other. This shift should be seen most clearly for female stereotype role nouns, as previous research revealed opposite representations for these, namely male-dominant for French and female-dominant for English.

Third, we predicted the shift to be modulated by speakers' L2 proficiency. We anticipated that it would be particularly apparent in the female stereotype condition, again, as French and English have been shown to generate two opposite representations in this condition.

3.3 Method

3.3.1 Participants

3.3.1.1 French-speaking sample

Sixty-one students from the University of Fribourg (Switzerland) took part in the experiment. They were all native French speakers (mean age: 22 years, range: 18-33; mean start age of L2 acquisition: 12 years, range: 3-12 years; mean years of L2 study: 7, range: 4-22 years; 44 female, 17 male). One participant was removed from the analyses, as their L2 proficiency was too low (less than a third correct on the C-test). Participants were granted course credits for experiment participation.

3.3.1.2 English-speaking sample

Sixty-six students from the University of Sussex (England) took part in the experiment. They were all native English speakers (mean age: 21 years, range: 18– 29 years; mean start age of L2 acquisition: 10 years, range: 3-19 years; mean years of L2 study: 10, range: 4-22 years; 50 female, 11 male). Five participants were removed from the analyses, as their L2 proficiency was too low (less than a third correct on the C- test). Participants were either granted course credits for experiment participation or paid £5.

3.3.2 Materials and design

3.3.2.1 Sensibility Judgment Task

All experimental and filler items were taken from Gygax et al. (2008). All role nouns had been tested for stereotypicality (Gabriel et al., 2008), likeliness of occurrence, and interpretation coherence. Each text was comprised of two sentence pairs in which the first sentence introduced a *female* (e.g., *social workers*_{English}, *assistants sociaux*_{French}), a *male* (e.g., *surgeon*, *chirurgiens*) or a *neutral* (e.g., *musicians*_{English}, *musiciens*_{French}) stereotypical role noun in the plural form in the English version and the masculine plural form in the French version as its main subject referent (see example [1a]).¹⁰ In French, the masculine plural form could be intended as generic, yet at the same time could have a possible masculine-only interpretation (see example [2a]).

The first sentence was then followed by a second sentence that mentioned a group of either men (see examples [1b] and [2]) or women (see examples [1c] and [2c]) referring to the group of people alluded by the role noun in the preceding sentence.

(1a) The social workers were walking through the station.

(1b) At the end of the day the majority of the men seemed to want to go home.

(1c) At the end of the day the majority of the women seemed to want to go home.

¹⁰To ensure that the participants would be familiar with all role nouns in the L2, we ran a pilot on 23 French–English bilingual speakers in which they had to translate the 36 experimental role nouns into their L1. Their L2 proficiency was measured on a self-assessment questionnaire. As no particular role noun seemed incomprehensible (for our French sample), with an average correct score of 82% of role nouns being familiar to them, we decided to keep all role nouns in the experiment and in the analyses.

- (2a) Les assistants sociaux marchaient dans la gare.
- (2b) A la fin de la journée plupart des hommes semblaient vouloir partir.
- (2c) A la fin de la journée plupart des femmes semblaient vouloir partir.

The sentence continuation including *men / hommes* and *women / femmes* would either match or mismatch the typicality of interpretation of the stereotypical role noun indicated in the first sentence. The neutral role nouns were the only role noun type that was intended to maintain an unbiased response, hence determined each language's general tendency.

Given that the findings reported by Gygax et al. (2008) were stable and generalizable across both participants and items (i.e., *F1* and *F2*), we divided the role nouns into two groups of equal stereotypicality distribution. Each participant read half of the experimental role nouns in French and the other half in English. For each language, there were six stereotypically female, six stereotypically male, and six stereotypically neutral role nouns, hence a total of 36 experimental items. In order to replicate Gygax et al.'s (2008) study in the context of our study, we constructed a total of four lists (two in each language) to ensure that each role noun was equally followed by a "men" or a "women" continuation in each language. If in one list a role noun written in French was followed by a male continuation, in another list, it would be followed by a female continuation, and in the two remaining lists, it would be written in English. Each participant read only one list. Creating these four lists allowed us to test participants in both languages without a repeated presentation of each role noun, which may have resulted in some confounding (repetition) effects.

Half of the participants began the judgment task in English and eventually switched to French, while the other half began with French and eventually switched to English; in other words, half of each group began the task in their L1 while the other half began with their L2, counterbalancing a possible effect of language dominance upon which the task began with. All experimental items were intended to elicit a positive "yes" response.

Thirty-six filler items – 18 in each language – were included to elicit a clear "no" response. There were three versions of filler items. One where there was a mismatch in the referents of the first and second sentence (see examples [3a, b] below), a second version where there was a mismatch in sex of the role noun mentioned in the first and second sentences (see examples [4a, b]), and finally pairs in which there was a semantic incoherence (see examples [5a, b]).

- (3a) The nannies were waiting on a bench.
- (3b) Because of the cloudy weather one of the graphic designers wore a raincoat.
- (4a) The chambermaids were crossing the hall.
- (4b) Due to the bad weather the majority of the men wore a raincoat.
- (5a) The florists were waiting in the rain.
- (5c) Since sunny weather was forecast some of the men weren't wearing a coat.

3.3.2.2 C-test

We chose to use the C-test to measure L2 proficiency as the C-test has been extensively researched in the field of language testing and has shown that it is a highly reliable source of an objective language proficiency measurement (Eckes & Grotjahn, 2006; Grotjahn, Klein-Braley, & Raatz, 2002; Klein-Braley & Raatz, 1984). The C-test is a form of a cloze test in which participants fill in the missing blanks formulated within a text (see Appendix for an example passage in English). The number of correct restorations indicates an overall efficiency of language processing, as it requires formation and anticipation of certain linguistic and grammatical constructions (Grotjahn et al., 2002). Each participant was asked to complete a C-test in his or her L2. For French native speakers, we selected and modified four of the five texts from the English C-test by Rahimi and Saadat (2005) and for English native speakers, a French C-test taken from Coleman (1994) was used. Each C-test consisted of four texts.

3.3.2.3 Self-evaluation questionnaire

Participants were asked for information of their L2 background regarding age of L2 acquisition, years of L2 study, and individual assessments regarding listening, reading, writing and speaking in their L2.

3.3.3 Apparatus

The experiment was conducted using an iMac for native English participants and a Power Macintosh 4400 for native French participants. It was controlled using the PsyScope Software (Cohen, MacWhinney, Flatt, & Provost, 1993) connected to a button box with two buttons labeled “yes” and “no”. Each item was presented on the computer screen and the “yes” button was controlled to maintain the handedness for each participant (i.e. the “yes” button would be the button pressed by the dominant hand).

3.3.4 Procedure

Each participant was tested individually in a small quiet room. All instructions were given in their respective native language. Participants were asked to read the sentences displayed on the computer screen in front of them and to judge whether the second sentence was “a sensible continuation of the first one”. In French, the expression *continuation possible* was used as in Gygax et al. (2008). Response to this question indicates the ease with which the second sentence is mapped onto the representation of the first one. The instructions stressed that participants should read at normal speed as they would normally read a book, and make judgments without prolonged contemplation.

Each trial began with a prompt (i.e., **** Ready? **** , **** Prêt? ****) of 250 milliseconds, subsequently followed by the first sentence. After reading the first sentence, a button press caused the second sentence (i.e., target sentence) to appear. Participants had to make a prompt decision of its sensibility by pressing either the “yes” or the “no” button. Participants were asked to keep the pointer of their dominant hand on the “yes” button and the non-dominant hand on the “no” button at all times during the experiment.

There were six practice trials in each participant’s L1 to familiarize them with the procedures of the experiment. After completing the sensibility judgment task, participants were given 20 minutes to complete the C-test in their L2. They were instructed to fill in as many blanks as possible within 20 minutes. Finally, participants answered the self-evaluation questionnaire.

3.4 Results

For each L1 language group, analyses were conducted on the proportion of positive responses and positive response times (in milliseconds) for the sensibility judgment tasks. All responses and response times were subject to mixed ANOVAs, treating participants as random factor. Analyses for items as random factors were not conducted, as our primary interest was within individual factors, specifically proficiency levels of each participant.

To account for individual differences and sentence length in response times, residual response times were calculated by fitting a regression equation of time against the number of characters in the second target sentence for each participant in each language. In L2, items in which the first sentence’s reading times were 2.5 standard deviations away from each participant’s mean were removed from the analysis for each language (2% of the data). These

longer reading times were considered as indicators of participants' struggle to understand the content of the sentence. Note that we do not present separate results for the two language orders (i.e., L1 or L2 first), as they did not show different patterns.

As noted earlier, the present study was grounded in three hypotheses. First we expected a male-dominant representation in French for French participants and a stereotype-dominant representation in English for English participants (Hypothesis 1). Second, we expected that, when changing language, participants' responses would signal a shift of gender representation, mostly observable in the female stereotype condition (Hypothesis 2). Third, we predicted that representation shifts would be modulated by participants' L2 proficiency, the effect of which being particularly apparent in the female stereotype condition (Hypothesis 3). More specifically, we expected French–English bilinguals whose L2 English proficiency is high to show a comprehension tendency in L2 English that resembles that of native English speakers, whereby “men” continuations will be favored over “women” continuations following stereotypically male role nouns and “women” continuations will be favored over “men” continuations after stereotypically female role nouns. On the other hand, less proficient French-English speakers were expected to present a preference for “men” continuations in L2 English regardless of the gender stereotypicality of the preceding role noun.

The latter hypothesis is based on the idea that L2 lexical representations of highly proficient speakers have stronger associations to semantic representations and consequently are less likely to be affected by indirect L2 to L1 lexical associations. As English role nouns are not grammatically marked for gender, their representations in L2 English should rely on gender stereotypicality. In contrast, less proficient speakers' access to semantic representations should be less direct (i.e., accessed through L2-to-L1 translation). As French role nouns are morphologically marked and associated with the male gender, low proficient speakers of L2 English, by accessing their L2 via L1 French, should be biased toward male representations. Given that English and French generate completely opposite representations when role nouns are female stereotyped, the effect of proficiency should be particularly apparent in this condition.

In the case of English–French bilinguals, highly proficient bilinguals should show in L2 French a preference for “men” continuations over “women” continuations regardless of the gender stereotypicality of the preceding role noun. Low proficient speakers' representations should rely on gender stereotypicality, as they should access their L2 via L1 English.

In order to test for these effects, for both the proportion of positive responses and positive response times, we first conducted an overall 2 (Proficiency: Advanced vs. Intermediate) $\times 2$ (Language: L1 vs. L2) $\times 3$ (Stereotype: Female vs. Male vs. Neutral) $\times 2$ (Continuation: Men vs. Women) mixed ANOVA, with Proficiency as a between-participant factor and Language, Stereotype and Continuation as within-participant factors. As we expected possible shift and proficiency effects to be seen most clearly for female stereotyped role nouns, we then ran the ANOVAs for female stereotyped role nouns only.

3.4.1 Proficiency measures

C-test scores were taken into account to establish L2 proficiency.¹¹ To split our sample into meaningful groups, we conducted hierarchical clusters using Ward's Method. This method enabled us to separate the participants into two meaningful groups (i.e., Advanced vs. Intermediate) without having to manually split our data (e.g., median split). All cluster analyses are briefly presented before the actual experimental results.

3.4.2 L1 French speakers

The cluster analysis on the C-test revealed two relatively even groups: an advanced group ($n = 35$) with a mean of 82.31 ($SD = 8.93$) and an intermediate group ($n = 25$) with a mean of 55.20 ($SD = 7.44$). The two groups were significantly different ($W_s = 325$, $z = -6.26$, $p < .001$).

3.4.2.1 Proportion of positive responses

The overall analysis revealed several significant effects: A Continuation effect, $F(1, 58) = 68.84$, $p < .001$, the proportion of positive responses being higher to “men” continuations (.76) than to “women” continuations (.54); a Stereotype effect, $F(2, 116) = 4.77$, $p < .01$, the proportion of positive responses to neutral stereotyped role nouns (.68) being significantly higher ($p < .05$) than to female stereotyped role nouns (.62), responses to male stereotyped role nouns being in between (.65); and a Proficiency effect, $F(1, 58) = 5.75$, $p < .05$, advanced participants giving more positive responses (.71) than intermediate participants (.60).

¹¹ Self-assessment scores of L2 proficiency were significantly correlated to the performance on C-tests scores, $r = .78$, $p < .001$.

Most importantly and as expected (Hypothesis 1), the source of the Continuation effect lied principally in the French part, as signaled by a significant Language \times Continuation effect, $F(1, 58) = 6.80, p < .01$ (see Figure 3.1). The difference in the proportion of positive responses to “men” and “women” was smaller in L2 English (.16) than in L1 French (.27). This interaction effect also supports the idea that when changing language, participants’ representation of gender shifts (Hypothesis 2).

There was also a Stereotype by Continuation effect, $F(2, 116) = 14.10, p < .01$, suggesting that the male bias was stronger in the male (.35; $p < .05$, with Bonferroni correction) and neutral (.18; $p < .05$) than in the female (.09; *ns.*) stereotyped condition. In essence, this is not surprising, as in the male stereotyped condition a male bias is fed in French by the masculine form just as in English by stereotypicality.

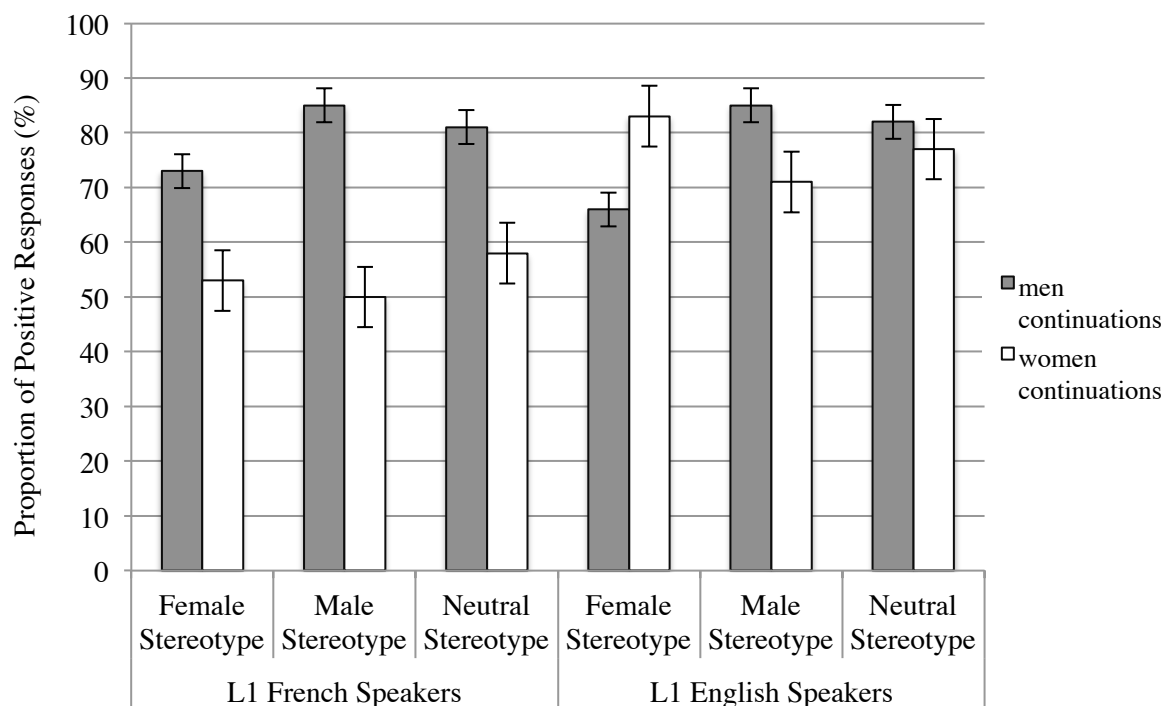


Figure 3.1 Proportion of positive responses of each native group in their respective L1 (error bars indicate standard error of the mean).

Finally, there was a trend toward a significant four-way interaction of Language \times Stereotype \times Continuation \times Proficiency, $F(2, 116) = 2.44, p = .09$, qualified, as expected (Hypothesis 3), by a significant Language \times Stereotype \times Continuation effect only in the advanced participant group, $F(2, 68) = 5.71, p < .025$ (intermediate group: $F(2, 48) = .66; ns.$).

To further examine this effect of Proficiency, and following our specific hypothesis on female stereotyped role nouns, we performed a planned 2 (Proficiency: Advanced vs. Intermediate) \times 2 (Language: English vs. French) \times 2 (Continuation: Men vs. Women) ANOVA only for responses to female stereotypical role nouns. As predicted, there was a significant Proficiency \times Language \times Continuation interaction, $F(1, 58) = 4.77$, $p = .05$, suggesting that if advanced participants shifted from a preference over “men” in L1 French (.75 for “men” and .52 for “women”) to a preference over “women” in L2 English (.63 for “men” and .73 for “women”), intermediate participants maintained a preference for “men” both in L1 French (.67 or “men” and .49 for “women”) and L2 (.64 for “men” and .52 for “women”) (see Figure 3.2).

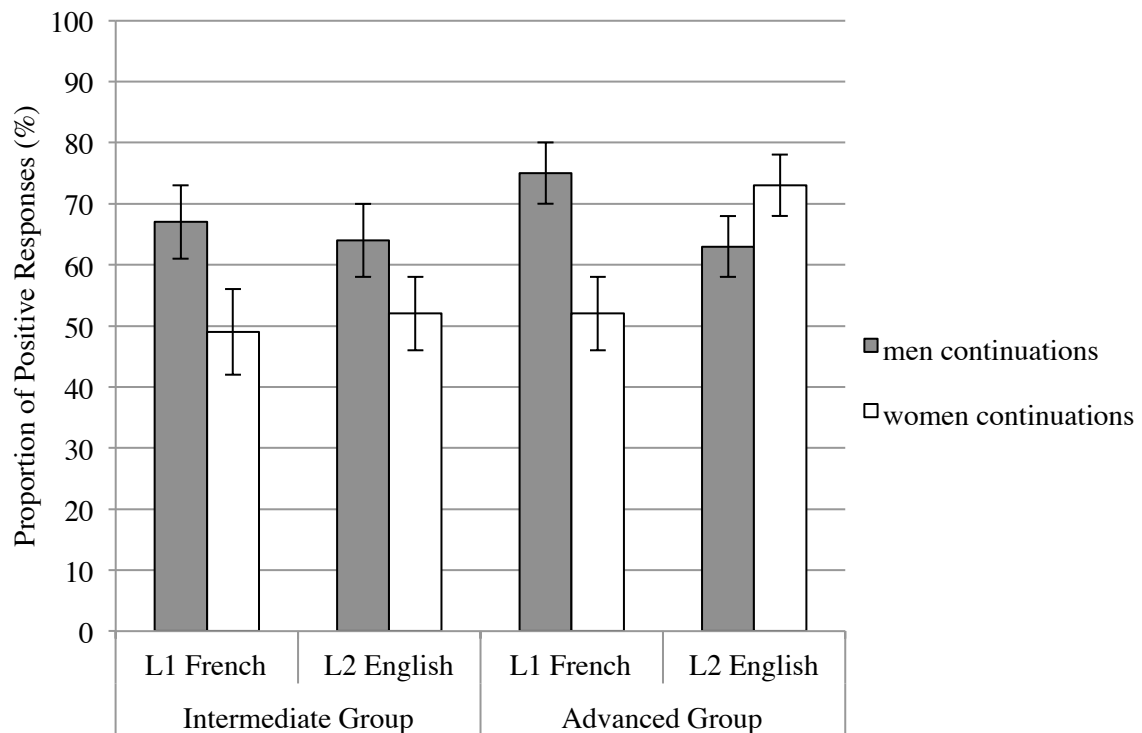


Figure 3.2 L1 French speakers’ proportion of positive responses of each continuation, only in the female stereotype condition. Proficiency is divided into intermediate and advanced groups (error bars indicate standard error of the mean).

3.4.2.2 Positive response times

There was only a significant Language \times Continuation interaction, $F(1, 58) = 7.06$, $p < .05$, showing that in L1 French, participants were 226 milliseconds faster to respond to “men” than to “women” ($p < .025$) continuations, whereas in L2 English, they were equally fast to

respond to “women” and “men” continuations (*ns*). This result also supports Hypotheses 1 and 2. No other main or interaction effects were significant.

3.4.3 L1 English speakers

The cluster analysis on the C-test revealed two relatively even groups: an Advanced group ($n = 40$) with a mean of 84.95 ($SD = 7.24$) and an Intermediate group ($n = 21$) with a mean of 54.95 ($SD = 9.58$). The two groups were significantly different ($W_s = 231$, $z = -6.38$, $p < .001$).

3.4.3.1 Proportion of positive responses

The overall analysis revealed a main effect of Language, $F(1, 59) = 5.52$, $p < .05$, participants responding more positively in English (.76) than in French (.71), and a Stereotype \times Continuation effect, $F(2, 118) = 12.62$, $p < .001$. This interaction was qualified by responses to “women” continuations (.78) being higher than to “men” (.68) continuations following female stereotyped role nouns, higher for “men” (.81) than “women” (.68) continuations after male stereotyped role nouns, and almost equal between the continuations following neutral stereotype role nouns (“women”: .70, “men”: .76) (*ns*).

The results also showed a crucial Language \times Stereotype \times Continuation interaction effect (see Figure 3.1), $F(2, 118) = 4.05$, $p < .05$, which confirmed that in L1 English, there was a Stereotype \times Continuation effect (Hypothesis 1), $F(2, 118) = 15.11$, $p < .05$, but in L2 French, there was not, $F(2, 118) = 2.41$, *ns*. In L1 English, both neutral and male role nouns were followed by a higher proportion of positive responses to “men” continuations than to “women” continuations (with the highest difference in the male stereotype condition), but, as expected, a higher proportion of positive responses to “women” continuations than to “men” continuations for stereotypically female role nouns. In L2 French, the effects were different, with “women” continuations receiving less positive responses for stereotypical female role nouns than in L1 English whereas “men” continuations receiving more positive responses, hinting at a shift toward a male-dominant representation in French, irrelevant of stereotype (Hypothesis 2). No other main or interaction effects were significant.

Following our specific hypothesis on female stereotyped role nouns and the effect of proficiency, we performed a planned 2 (Proficiency: Advanced vs. Intermediate) \times 2 (Language: English vs. French) \times 2 (Continuation: Men vs. Women) ANOVA only for the

female stereotypical role nouns. Contrary to our expectations, no interaction effect with Proficiency was found, invalidating Hypothesis 3 (see Figure 3.3).

3.4.3.2 Positive response times

The results revealed a Continuation effect, $F(1, 59) = 6.09, p < .05$, participants responding 275 milliseconds faster to “men” continuations than to “women” continuations, which was further qualified by a significant Language \times Continuation interaction, $F(1, 59) = 8.09, p < .01$. Participants responded equally fast to “men” and “women” continuations in their L1 English but 642 milliseconds faster to “men” continuations than to “women” continuations in L2 French ($p < .025$), suggesting a greater male bias when reading in French than in English (Hypothesis 2). No other main or interaction effects were significant.

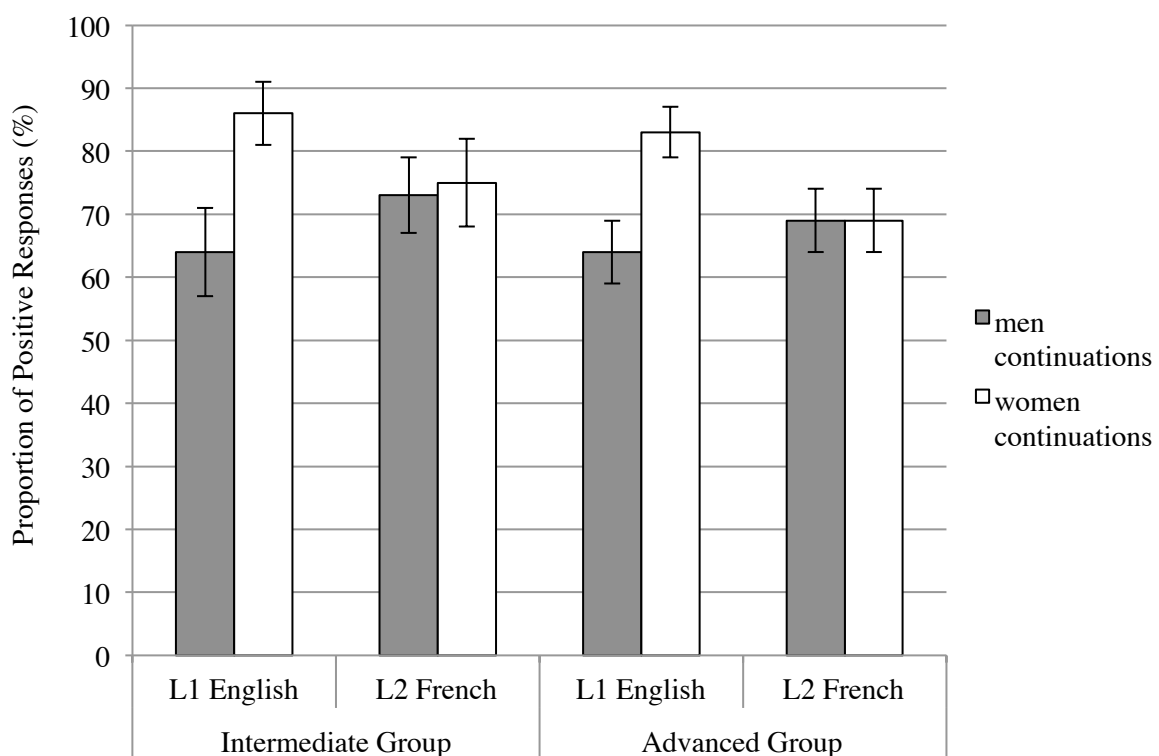


Figure 3.3 L1 English speakers’ proportion of positive responses of each continuation, only in the female stereotype condition. Proficiency is divided into intermediate and advanced groups (error bars indicate standard error of the mean).

Again, we tested our specific hypothesis on the effects of Proficiency on female stereotypical role nouns. Though the analyses revealed a significant Language \times Continuation interaction, $F(1,59) = 4.47, p < .05$, revealing that in L1 English, participants responded 235 milliseconds faster to “women” continuations than to “men” continuations for female

stereotypical role nouns ($p < .025$) whereas responses to “men” and “women” continuations did not differ in their L2 French (*ns.*), and there was no effect of proficiency.

3.5 Discussion and Conclusion

The present study investigated the influence of grammatical gender and stereotypical information on gender representation in bilinguals of different L2 proficiency that speak both a grammatical gender language (French) and a natural gender language (English).

Results showed that the differences in the comprehension patterns that were previously reported by Gygax et al. (2008) proved to be robust in each language group’s native language. In English (L1), when participants encountered a stereotypical role noun, they relied on stereotypical information to make inferences regarding the referent’s sex whereas in French (L1) they were prone to rely on the specificity of the masculine form (i.e., masculine form = male) even if stereotypical information was readily available to them (Hypothesis 1).

As we believed that L1 should have been impervious to the influence of L2 given the complexity of acquiring and adjusting new grammatical systems – unless newly formed L2 grammatical gender structures are firmly established – we did not predict (and did not observe) an effect of reverse transfer. In other words, transfer of L2 features onto L1 (Brown & Gullberg, 2008; Jarvis & Pavlenko, 2008) was neither expected nor observed. However, in the previous study that we compare our results to (Gygax et al., 2008), participants’ L2 (or L3) proficiency was not reported. Thus, we cannot know if previous samples were built on monolingual or bilingual participants and as a consequence we cannot definitely exclude the presence of a reverse effect. In this regard, we cannot directly compare our results to the initial study and hence, in order to rule out possible effects of L2 transfer onto L1, future research may address this issue by complementing a control (truly) monolingual group. Note that an extensive comparison to previous samples is also made difficult by the constraining design employed in our experiment, whereby participants generated, in L1 (and this is not even considering the fact that we also split our sample into proficiency), only half of the data of those in the original study. Nevertheless, our L1 results are in line with previous findings suggesting that morphosyntactic cues, here grammatical gender, strongly influence the way gender inferences are generated in constructing representations of protagonists’ gender from text.

In L2, participants' responses seem to conform, at least partly, to those of the native group (Hypothesis 2). English participants showed a male-dominant comprehension tendency in French (mostly signaled by increased responses to female stereotype), and French participants showed a decrease in constructing male-dominant representations in English, signaling the reliance on stereotype as a source of information for making gender inferences.

These findings show that a switch of language when processing role nouns essentially alters the way readers mentally represent groups of people (i.e., in their mental models of the text), specifically in terms of gender. French-speaking readers generate more male-dominant inferences in L1, whereas they are prone to activate stereotypical information when switching to L2 English. On the other hand, English speakers move from stereotypical representations to male-dominant ones when switching from English to French.

In essence, when switching from a grammatical gender language to a non-gendered one (and vice versa), readers switch from one bias to another. Put differently, local morphological elements of each language appeared to have emphasized gender-based associations, which in turn influenced the comprehension of gender-associated information. When interpreting these findings, one should, however, note that our study was based on a linguistic task, and hence our study addresses issues associated with a semantic (in contrast to a conceptual) level of representation.

Our findings also revealed that even if in both languages there was a substantial shift in representation, the resulting representation did not fully match that of the native group. Studies converge on the idea that language proficiency appears to be an important indicator of the degree to which language affects cognition and L2 task performance, given that its complexity is constructed of multiple factors such as age, environment, motivation, affectivity, native language or years of study to say the least. However, given that no person's L2 learning experience is unique, the means to operationalize and define language proficiency differ among studies. For example, some studies adapt self-assessment measures using language history questionnaires completed by participants themselves (Malt & Sloman, 2003) or a combination of such questionnaires and other tasks such as lexical decision tasks in the L2 (Experiment 2 in Kroll, Michael, Tokowicz, & Dufour, 2002). Silverberg and Samuel (2004), who examined proficiency and age of L2 acquisition, combined language history questionnaires and the Boston Naming Test (Kaplan, Goodglass, Weintraub, & Goodglass, 1983), whereas Athanasopoulos and Kasai (2008) took into account participants' length of

stay in the L2-speaking country, their performance on the Oxford Quick Placement Test (QPT, 2001), and a picture description task concentrating on specific grammatical properties.

In the present study, we operationalized bilingual proficiency levels in terms of an objective evaluation criteria assessed by C-test performance, which has been shown to measure comprehensive language competence, and found that the linguistic competence measured by C-test scores was a good predictor of the influences of language onto gender representation. The extent to which these processing shifts were displayed differed in function to the comprehenders' L2 proficiency (Hypothesis 2).

The results from the proportion of responses for the less-proficient French-English bilinguals appeared to conform to the male-dominant representation tendency of their L1 French, showing a greater preference for “men” than “women” continuations in both female and male stereotypes, yet the preference for “men” continuations was not observed for female stereotypes for advanced participants (quite the contrary). This change in representation suggests a gradual shift to a stereotype bias modulated by bilinguals' increasing L2 proficiency (Hypothesis 3).

Our results from the English native speakers did not yield any effects of proficiency in terms of modulating the processing switch, rather the effect of proficiency was primarily seen among the French native speakers. We believe that the observed effect for proficiency in our data was mostly apparent among French participants inasmuch as our French native sample was taken from Switzerland, a multilingual country where the language context is more dynamic and English would often be used as a lingua franca on a day-to-day basis. This social context would most likely give the Swiss-French participants an advantage over English participants (although the proficiency scores were quite similar).

It could also be argued that this language difference is bound to a language shift complexity. Though the debate as to the extent to which speakers can fully acquire an L2 grammatical system has not been resolved, studies suggest that even among speakers of grammatical gender languages, the acquisition of a new gender system in another language is relatively difficult and that learners show persistent errors (Dewaele & Véronique, 2001; Franceschina, 2001). For example, in acquiring a language without a grammatical gender system, French speakers need to adjust their established grammatical system (i.e., French has also non-gendered role nouns such as *artiste* [artist, painter]), and English speakers need to create a new way of mapping gender per se. This modification is not simple for either language group, given that gender information encompasses both grammatical and conceptual

properties. In the context of our study, it is reasonable to assume that the interpretation of the masculine form in L2 French by native speakers of a language with no grammatical gender (i.e., English) is extremely intricate, as those speakers have to acquire additional grammatical particularities (also difficult for native French speakers). The fact that French native speakers have a more compound foundation of gender-associated information (i.e., grammar and stereotype) embedded in their system may give them an advantage to resolve such information in English where such features are less complex.

The intricate nature of grammatical gender in L2 French could also lead to a different explanation of the lack of proficiency of L2 French speakers. One could argue that grammatical gender information is overt, whereas stereotypical information is covert. In this sense, stereotypical information is elaborative and conceptual in nature, requiring (extensive) world knowledge, whereas gender grammatical cues may simply stand out as very different to one's L1, for both proficient and less proficient L2 French speakers. As a consequence, when accessing a noun's concept, if the lack of L2 proficiency should normally impose a passage through L1 (as for low proficient L2 English speakers in the female stereotype condition), the prominence of the rather unaccustomed masculine grammatical cue may bring intermediate and advanced proficient L2 French speakers closer in their reliance on grammatical cues. In all, the discussion therefore appears to extend beyond the issues of comprehension tendency of gender-associated information in each language but also speaks of different language switch complexity.

One may argue that some of our language switch effects were for at least some participants actually hampered by some of the cognate nouns between the inter-lingual items found in our study. This critique is plausible, given the numerous findings on the cognate facilitation effect in bilingual language production and (Costa, Caramazza, & Sebastian-Galles, 2000; Dijkstra et al., 1999). For example, the English noun *golfers* shares orthographic (and phonological) features with the French noun *golfeurs*. For both bilingual groups, reading L2 nouns that are cognates to their L1 counterparts may force access to semantic representations in L1. For French-English bilinguals, for example, reading *golfers* might activate the semantic representation of the French role noun *golfeurs* (cognate), associated with “men” as a result of its grammatical male feature. Likewise, for English-French bilinguals, reading the French role noun *golfeurs* might activate the semantic representation of *golfers*, which would not be associated with any male feature, resulting in a male-attenuated representation. To clarify this issue, we removed all possible cognates from

our data set and re-ran our analyses. In neither analyses did the results change as a function of removing cognates, further supporting our hypotheses that this magnitude in comprehension shift appeared to have been modulated by participants' proficiency levels.

In effect, our results are consistent with those reported by Athanasopoulos and colleagues (Athanasopoulos et al., 2011; Athanasopoulos & Kasai, 2008), in which they found bilinguals' cognitive behaviors in L2 to resemble a pattern that was "in-between" (Athanasopoulos et al., 2011, p. 14) the native speakers of the L1 and L2, and that the degree to which these patterns manifested depended on language proficiency (Athanasopoulos, 2009; Athanasopoulos et al., 2011; Athanasopoulos & Kasai, 2008).

Although modest, the effects of proficiency were apparent in our data, mainly supporting the idea that as proficiency increases, representations in L2 tend to mimic those of native speakers. In a sense, lower proficient speakers' L2 representations, being in-between the native ones, were less biased.

A final issue that should be addressed is the adapted paradigm of our study. The sensible continuation paradigm employed in our study has been commonly implemented in studies investigating the interpretation of gender on mental representations (Gabriel & Gygax, 2008; Garnham, Gabriel, Sarrasin, Gygax, & Oakhill, 2012; Gygax et al., 2008). The advantage of this task is that it is effective in addressing the ease with which certain types of information are integrated into ongoing mental models, and thus was appropriate for our primary goals in assessing the influences of language over semantic representations. Our data from each native group's L1 conformed to the monolingual data reported in Gygax et al. (2008) showing the robust nature of native language processing regardless of the possible influence of L2, whereas a gradual transfer of comprehension tendency modulated by L2 proficiency was observed in participants' non-native L2. It should be noted though that other bilingual studies on grammatical gender have also concentrated on the effects of language on conceptual representation, that is, on tasks that were non-linguistic per se (e.g., color categorization: Athanasopoulos et al., 2011; Athanasopoulos & Kasai, 2008; gender voice-attribution: Bassetti, 2007; Sera et al., 2002). Still, our results are in line with the *thinking-for-speaking* notion proposed by Slobin (1996a, 2000, 2003) stipulating that semantic characteristics of a language may influence language processes in another language. More concretely, certain language-specific patterns may direct and accentuate our attention to particular dispositions, such as events or categories, wherein speakers of one language may favor focusing their attention on one aspect and speakers of another language, on another.

This process constitutes one of the essential components of what Slobin refers to as “*language in use*” in which speakers systematically code experiences required by the language for subsequent language output and for the online construction of mental representations. Our data therefore highlight the notion that even small linguistic features can influence mental representations for the purposes of language comprehension.

To conclude, these results presented here alone are not sufficient to substantiate the direct influence of language over non-linguistic cognition (i.e., conceptual representation *per se*), but most likely support the claim made by Slobin (1996b, 2000, 2003), suggesting that certain morphosyntactic features of a language may emphasize certain linguistic aspects, hence influencing certain mental representations. These results hence suggest that language constitutes the driving force for influencing certain processing functions, rather than comprehenders’ processing dispositions. In other words, given that French seems to accentuate readers’ attention toward male frames of reference and English toward stereotypical conceptualizations, mental representations of gender created by bilinguals who speak languages with different gender features appear to alternate as a function of the language at use.

Chapter 4 : Study II

Representing the social face of grammatical and stereotypical gender: When gender-associated stereotypical information does count

The study presented in this chapter has been submitted for publication as:

Sato, S., Gyga, P., & Gabriel, U. (2013). Representing the social face of grammatical and stereotypical gender: When gender-associated stereotypical information does count.

4.1 Abstract

The findings from Study I revealed that bilinguals adhered to the gender representation tendency of the language in which information was being encoded. A switch from one language to the other indicated that their representations changed accordingly, with their L2 proficiency levels influencing the magnitude in which the language-bound regularities would manifest in the readers' L2. While in English, readers relied on stereotypical knowledge to represent gender, in French, readers constructed a male-dominant representation based on the role nouns presented in the masculine forms, suggesting that the surface forms semantically contributed to the representation of gender. As the effects of stereotype information during the comprehension of grammatical gender languages was overridden by grammatical gender information in the previous study, the present study aimed to further investigate this impact by employing a more sensible experimental paradigm.

In particular, this study examined the interaction of stereotypical and grammatical information when representing gender stereotypical role nouns in French and German. In a priming task employing a combined linguistic-visual paradigm, bilingual speakers completed experimental trials in which they were first shown a gender stereotypical noun in the masculine plural form in German or French, and then were asked to judge whether a pair of facial images representing *two men* or *a man and a woman* could represent the prime. Differences in representations associated to the German and French plural determiners *die* and *les* [the] were also examined. Results suggested that although gender-associated stereotypical information is spontaneously activated at a lexical level, the representation is still dominated

by the overt masculine grammatical form (though interpretable as *generic*) of the role nouns. The results further indicated that the employed linguistic-visual paradigm was well-suited to gauge the impact of stereotype information, the latter often lacking scrutiny in past research that adapt linguistic-linguistic paradigms. No effect of the determiners was found, further confirming the importance of the grammatical structure of the role nouns. The results provide further support for *thinking-for-speaking* effects (1996a, 2000, 2003) suggesting that linguistic encoding of gender contributes to shaping our gender concepts during language comprehension.

4.2 Introduction

Psycholinguistic research investigating the mechanisms underlying the mental representation of gender in grammatical gender languages such as French and German have examined its core principals applying different linguistic paradigms. Among them, prominent were those taking into consideration behavioral measures such as reading times (e.g., Carreiras, Garnham, Oakhill, & Cain, 1996; Irmen & Roßberg, 2004), eye-fixation measures (e.g., Esaulova, Reali, & von Stockhausen, 2013; Irmen, 2007) and response proportions (Garnham, Gabriel, Sarrasin, Gygax, & Oakhill, 2012; Gygax, Gabriel, Sarrasin, Oakhill, & Garnham, 2008; Sato, Gygax, & Gabriel, 2013). A commonality among these paradigms is the procedure by which participants are instructed to process experimental stimuli: a prime intended to activate a probabilistic conceptual gender (e.g., *politicians* activates the male concept) is followed by a target stimulus that includes a reference to a definitional gender (e.g. *elle* [she], *plusieurs hommes* [several men] in French). According to the conceptual (e.g., stereotype) and / or grammatical congruency between the prime and target gender (e.g. *the nurse* - *he*), the different paradigms are expected to reveal facilitation or difficulty in integrating the information. Measurements associated to gender match-mismatch effects are interpreted as reflecting the intricate interaction between grammatical and stereotypical gender information.

A large amount of data from these paradigms attest that in grammatical gender languages, there is a general preference to represent the male gender associated to the default usage of the masculine grammatical form (e.g., Gygax et al., 2008; Irmen, 2007; Lévy, Gygax, & Gabriel, 2014). In languages with no grammatical gender (e.g., English), gender stereotype information constitutes the basis for gender representations (Oakhill, Garnham, & Reynolds, 2005; Reynolds, Garnham, & Oakhill, 2006). Although research with these

paradigms has produced significant knowledge, it remains open whether the dominance of male representations in grammatical gender languages may well be prompted by the very nature of the prime and target stimuli, meaning that both are *verbal* stimuli. The use of verbal target stimuli, maintaining a close link with its verbal prime, may result in mental representations that only (and merely) reflect linguistic activations. Namely, we argue that processing both prime and target stimuli in a verbal context may compel readers to over-monitor grammatical and syntactical properties. This monitoring in turn may enhance the signal of a representation based on linguistic cues (i.e., toward a male bias in gender-marked languages). In the study presented here, we approach these methodological issues in gender representation research by applying visual (i.e., facial images) rather than verbal target stimuli. Rather than presenting texts that require readers to link textual elements, a lexical priming approach was preferred to directly examine lexical gender activation (as suggested by Oakhill et al., 2005). Bearing in mind these issues, we expected to provide a more direct and conceptual (as opposed to linguistic) examination of how readers construct a representation of gender when processing grammatically marked role nouns.

More generally, we address the idea that the way concepts are linguistically encoded in a language may shape and influence our mental representations. Methodologically, we chose bilinguals as an experimental sample as they allow us to examine the extent in which mental representations are directly affected by linguistic regularities. We hypothesize that bilinguals' representations should bear different characteristics in each language according to their linguistic features. Representations that bear the characteristics of each language provide evidence for the impact of language effects.

In what follows, we first discuss the relevance of language, as impacting gender representations in general. We then more specifically present how grammaticization of gender information influences gender representations, which will set the theoretical ground for our study.

4.2.1 The impact of language on the construction of gender representations

The idea that specific languages and the ways in which concepts are organized in their linguistic structures shape our interpretation of verbal stimuli has long been of interest in linguistics and psychology (Gennari, Sloman, Malt, & Fitch, 2002; Papafragou, Massey, & Gleitman, 2002). Further developed as the *thinking-for-speaking* hypothesis by Slobin

(1996b) in his work on motion events, the notion proposes that the encoding of concepts and events within a language acts both as a foundational and as a constraining structure for how verbal information is represented. In other words, processing a specific language imposes speakers to focus on particular concepts that are lexicalized and / or grammaticized within its structure, resulting in language-bound representations. As will be further discussed in this paper, bilinguals are particularly suited for testing the thinking-for-speaking hypothesis as they offer an interesting platform to examine the extent to which comprehension mechanisms change as a function of the characteristics of the language being used, or whether representation tendencies abide with the comprehension mechanisms that the language user is accustomed to (i.e., first language [L1]) (e.g., Boroditsky, Schmidt, & Phillips, 2003; Bylund & Jarvis, 2011; Fausey, Long, Inamori, & Boroditsky, 2010). In terms of gender representations, recent research suggests that the ways in which gender information is encoded may offer an attractive explanation for the diverging gender representations observed between typologically distinctive languages (e.g., Sato et al., 2013). For example, when readers read about person reference role nouns in English (e.g. *the mechanic*), they generate inferences about the possible gender of the depicted person based on their knowledge of the world, gender stereotypes acting as a primary source of information (Kennison & Trofe, 2003; Kreiner, Sturt, & Garrod, 2008). In a seminal priming study by Banaji and Hardin (1996) participants were presented with either a stereotypical (e.g., female stereotype: *nurse*, male stereotype: *mechanic*, neutral stereotype: *reporter*) or gender definitional (e.g., female definitional: *mother*, male definitional: *king*) prime followed by a pronoun target. The authors found that participants' responses to judge if the target stimuli was *male or female* (Experiment 1: *he vs. hers*) or a *pronoun or not* (Experiment 2: *she vs. do*) were faster when there was a gender congruency between the prime and target stimuli both when following stereotypical and gender definitional primes. Although these results were interpreted as being demonstrative of an automatic activation of gender inferences, the study was furthered by Oakhill et al. (2005). In their study, participants were required to make more conscious decisions regarding the link between the role noun prime and the gender-definitional kinship target (*sister, uncle*) in a task to judge if the two words could refer to the same person. Comprehension facilitation was examined with shorter response times being observed when the stereotypical gender was congruent with the gender of the kinship term, showing that a specific gender inference associated to existing stereotypes had been activated. This effect in activating stereotyped gender inferences has been found to be immediate and robust among

English readers, showing that such role nouns may prime a specific stereotypical gender even if morphological or grammatical information does not require readers to do so.

These results however, are not readily generalizable for readers of grammatical gender languages such as French or German, where stereotypical gender is only one of two sources contributing to the construction of gender representations. In these languages, gender is also integrated as part of the grammatical structure of the language, with grammatical gender determining a specific gender category to all nouns (e.g., masculine, feminine, neuter) and the grammatical gender of most person reference role nouns corresponding to the biological gender of the referent (i.e., masculine = man, feminine = woman). In such a case, grammatical gender unambiguously identifies a role noun such as *nurses* as being either a woman (*infirmières*_{Feminine} [female nurses]) or a man (*infirmiers*_{Masculine} [male nurses]) according to its grammatical marking (see Stahlberg, Braun, Irmen, & Sczesny, 2007, for a further overview on classifications of languages based on gender information). This linguistic feature within the language structure requires that its language users consistently monitor gender information at both grammatical and semantic levels. A fundamental claim made by researchers is that the interaction of stereotypical and grammatical gender information during the processing of role nouns is complex, and the mechanisms for representing gender information are not always straightforward (e.g., Esaulova et al., 2013; Garnham et al., 2012; Gygax, Gabriel, Lévy, Pool, Grivel, & Pedrazzini, 2012; Irmen, 2007).

This complexity is rendered by the semantic associations linked to grammatical gender. Whereas role nouns such as *infirmières*_{Feminine} marked in the feminine grammatical form refer unambiguously to female nurses, there is a discrepancy between form and meaning for the masculine form (i.e., *infirmiers*_{Masculine}). Either it can refer exclusively to men (i.e., only male nurses) or it may refer to a group composed of both male and female persons (i.e. *generic* interpretation). While readers are presented with a challenge to disambiguate the intended interpretation of the masculine form, its surface form naturally emphasizes the association to the male gender, inevitably prompting a male-specific interpretation. Adapting the priming task used by Oakhill et al. (2005), Gygax et al. (2012) illustrated a strong male-specific activation triggered by the masculine grammatical form. In their study, French-speaking participants were instructed to decide whether the person represented by a kinship term in pairs such as *tante* [aunt] – *infirmiers*_{Masculine} [nurses] could belong to a group represented by the second noun (always in the grammatical masculine plural form). Participants responded more often positively (i.e., yes) and faster when the kinship term was a

man, indicating a male dominant representation. Importantly, this male dominance persisted irrespective of the stereotype of the role noun. When participants were reminded half way through the experiment that the masculine form could encompass a generic interpretation, although more positive responses were given when female kinships were presented, response times did not change. The authors concluded that the generic interpretation could only be activated through active processes, yet the male-specific interpretation was always passively activated. Most studies using on-line (e.g., Gabriel & Gygax, 2008; Gygax et al., 2008) and off-line (Braun, Szczesny, & Stahlberg, 2005; Stahlberg, Szczesny, & Braun, 2001) tasks concur on the male-specific impact of the masculine form.

Another issue with regard to the interpretation difficulty of masculine forms is that role nouns such as *infirmiers*_{Masculine} [nurses] may convey stereotype information that does not necessarily correspond to the gender of the grammatical form. While researchers using lexical-based paradigms have often demonstrated the facilitation to activate the male-specific interpretation of the masculine grammatical form, stereotype influences in grammatical gender languages have been found only in particular circumstances. For example, Esaulova et al. (2014) found that anaphor resolution, whereby an anaphor (e.g., he) is linked to an antecedent (e.g., the mechanics), was influenced by gender stereotypicality in German, but only in late temporal eye-tracking measures. Additionally, Cacciari and Padovani (2007), using the same experimental framework as Banaji and Hardin (1996) in a grammatical gender language, found stereotype effects in Italian, but only when testing Italian *bi-gender* role nouns, which are not grammatically marked for gender (e.g., female stereotype: *insegnante* [teacher]).

In sum, most studies have shown a strong impact of grammatical gender, yet some authors claim that grammatical gender had only overshadowed stereotype effects. Although the impact of grammatical cues seem central in representation processes, the reasons for their dominance over stereotype information are not yet clear. We approach this impact of grammaticization of information on mental representations through a theoretical context offered by the thinking-for-speaking hypothesis.

4.2.2 The impact of grammaticization of information on mental representations across languages

The importance of grammatical gender discussed so far demonstrates how grammaticized information influences readers' comprehension processes. Grammatical

gender languages impose their users to consistently monitor gender both on grammatical and semantic levels. If as suggested by the thinking-for-speaking hypothesis, information grammaticized in languages shapes readers' semantic concepts, which in turn anchor their representations, these regularities should also surface on their representations. Alternatively, it could be the case that other languages (and their structures) generate their own biases, even for readers who switch from one language to another. In other terms, readers of more than one language may switch representations as they switch language. This notion is further developed in this study by looking particularly at bilinguals where the language biases of each of the bilingual's languages should become evident on their representations. Following this line of reasoning, Sato et al. (2013) investigated, in a sentence-based paradigm, whether English-French bilinguals would construct different representations according to their L1 and L2 (second language). They presented English and French bilingual participants with sentence primes including role nouns with stereotypical gender (e.g., female: *nurses*, male: *politicians*, neutral: *pedestrians*). Participants judged the plausibility of target sentences including a gender reference (e.g., *some men*, *some women*) to be a sensible continuation of the prime. The results indicated that switching language was also accompanied by changes of biases in mental representations of gender, with English eliciting stereotyped representations and French male-biased representations triggered by the masculine form. Importantly, participants' L2 proficiency, as measured by a C-test (e.g., Eckes & Grotjahn, 2006; Klein-Braley & Raatz, 1984) was a good indicator of the extent of the representation switch between L1 and L2.

In the present study, we focus particularly on French and German as opposed to English for the following two reasons. First of all, English as examined in Sato et al.'s (2013) study, is arguably an unsatisfactory candidate for testing thinking-for-speaking effects, as it lacks grammatical gender. In contrast, French and German make for better candidates as gender is systematically structured within their grammatical systems. Thus, characteristics surfacing on representations in French and German should reflect purely linguistic effects and their impact of how linguistic encoding contributes in shaping gender representations.

Second, despite their shared status as grammatical gender languages and their common usage of the masculine form to denote a generic interpretation, gender associations linked to their plural determiners differ in the two languages. Although in French the plural determiner *les* [the - plural] is gender neutral, the German equivalent *die* [the - plural] shares the same surface structure as the singular feminine determiner *die* [the – singular – feminine].

This implies that while the French masculine role noun commonly appears in contexts presented with a gender-neutral determiner, the German role noun may be presented with a determiner that shares an association to the female gender. Rothermund (1998) explored this female association in his study investigating gender representations in German. In his study, participants were presented with short written scenarios that included either person references with a specified gender (e.g., *Herr Schmidt* [Mr. Schmidt] or *Frau Meiser* [Ms. Meiser]) or with a masculine intended as a generic phrase in the singular or plural form (e.g., *der Student* [the student] or *die Studenten* [the students]). In a recognition task that followed, participants had to either reject or accept test phrases that had specific gender associations. Although scenarios with a singular masculine reference increased rejection times for sentences with female associations indicating a male bias, scenarios with masculine plural references decreased rejection times for sentences with female associations, suggesting a female bias. The latter effect was unanticipated, yet was attributed as being triggered by the female equivalent determiner.

Similarly, Garnham et al. (2012) found a male attenuated effect (or an additive female effect) when presenting the German plural pronoun *sie* (i.e., *they* – also feminine-equivalent) in a sentence continuation paradigm. When the same was done in French however, the masculine pronoun *ils* [they – masculine *specific* or *generic*] did not have a male amplifying effect despite its male association. The authors argued that although cumulating male grammatical cues does not augment male biases, combinations of male and female-equivalent grammatical cues may distract readers from activating male specific representations. When looking strictly at determiners, only several studies (e.g. Gygax et al., 2008) have generated specific hypotheses as to the impact of the definite plural determiner *die* in German, yet its female-bias effect (as shown by Rothermund, 1998) was never clearly replicated. If in the present study we are able to observe differences in gender biases between French and German representations, such as a reduced male bias for German representations when role nouns with determiners are presented as primes, it should provide compelling evidence as to the direct impact the grammaticization of language has on our conceptualization of gender information.

4.2.3 The present study

In sum, the findings of past studies illustrate a complex mechanism at stake when processing gender information. When processing a stereotypical role noun in grammatical gender languages, surface-level cues of grammatical gender play a crucial role. If the

grammatical gender is feminine, the interpretation is straightforward given only one possible interpretation (i.e., female gender). If the grammatical marking is masculine, readers are faced with some level of ambiguity given possible alternative interpretations. In such a case, the facilitation for a male-specific interpretation is tangible given the stronger association to its surface form (Lévy et al., 2014).

Most studies indicate that the activation of the male-specific interpretation through surface level cues appears to be sufficiently strong to even override possibly co-activated stereotype information. If the female attenuating effects reported by Rothermund (1998) and Garnham et al. (2012) are reliable, they suggest interesting implications in terms of minimizing, or neutralizing male biases. In this regard, the comparison of French and German offers the possibility to investigate the subtle influences that surface forms play on shaping readers' gender representations.

When role nouns in the masculine forms are presented in conjunction with German determiners that bear associations to the female gender, any male biases triggered by grammatical gender may be attenuated. In contrast, such male-bias attenuation effects should not be observed in French given that its determiner lacks any gender association.

Combining both verbal and visual stimuli in a facial judgment task, this study investigates the effects of different linguistic encodings of gender and stereotypical gender on the representation of person reference role nouns. In particular, the interpretation of masculine role nouns and the immediacy of which different gender effects come into play were of central concern. While past studies have based their findings for gender representation on anaphor resolution paradigms, which relied on the detection of semantic and syntactic inconsistencies in comprehension, they do not strictly speak to the immediacy of the activation of such surface-level grammatical cues. In those studies, discursive contextual elements may interfere with stereotype activation or with the accessing of signals during activation. Note that only a handful of studies have been conducted on gender representation using a lexical-based paradigm (Banaji & Hardin, 1996; Cacciari & Padovani, 2007; Gygax et al., 2012; Oakhill et al., 2005; Siyanova-Chanturia, Pesciarelli, & Cacciari, 2012), although never directly addressing the impact of the use of the masculine form, or the impact of role noun determiners. We expect that the conceptual nature of the paradigm we propose, as well as its capacity to access any immediate activation signals should enable us to gauge the subtle stereotype effects that has lacked research scrutiny in studies using different paradigms.

In the task we devised, participants had to make judgments as to whether a visually presented pair of faces could represent a preceding role noun prime. As mentioned earlier, these visual target stimuli were intended to examine whether male biases found in previous studies were prompted by the use of verbal target stimuli, and not by the actual mental representation triggered by a role noun. The composition of face pairs represented the possible interpretations that the masculine form holds (i.e. a male specific and a generic interpretation). French and German were taken as comparative languages as they differ in the gender associated to the surface structures of their plural definite determiners, yet offering a common grammatical gender system. Role nouns were thus presented either in the presence or absence of definite determiners in each language. The experimental task was carried out in participants' L1 and L2 in order to examine any representational shift that would be prompted by the regularities of each language.

4.3 Method

4.3.1 Participants

4.3.1.1 German-speaking sample

Fifty German-speaking students from the University of Fribourg (Switzerland) participated in the experiment for course credits. All participants were native speakers of German whose L2 was French (mean age: 22, mean start age of French acquisition: 9.4 years, mean number of schooling of French as L2: 7.2 years). Forty-one participants were women¹².

4.3.1.2 French-speaking sample

Fifty-one French-speaking students from the University of Fribourg participated in the experiment for course credits. All participants were native speakers of French whose L2 was German (mean age: 22, mean start age of German acquisition: 7.5 years, mean number of schooling of German as L2: 9.2 years). Thirty-nine participants were women.

¹² As past studies on gender representation (e.g., Garnham et al., 2012; Gygax & Gabriel, 2008; Gygax, et al., 2008) did not find effects of participants' gender in reading tasks, we did not balance the gender sample of our participants.

4.3.2 Materials

4.3.2.1 Prime role nouns

Thirty-six gender stereotypical role nouns were selected as primes for the experiment. These role nouns were taken from Gygax et al. (2008), all of which were normed and tested for gender stereotypicality in Gabriel, Gygax, Sarasin, Garnham and Oakhill (2008) in both German and French. Role nouns were female (e.g., nurses [*Krankenpfleger / infirmiers*]), male (e.g., bosses [*Arbeitgeber / patrons*]) or neutral (e.g., pedestrians [*Spaziergänger / promeneurs*]) in stereotype. A list of the role noun used is presented in Table 1.1.

4.3.2.2 Target face pairs

The face pairs were created with the face modeling software FaceGen© Modeller program version 3.1.4 (Singular Inversions Inc., 2004). A total of 30 male and 30 female Caucasian faces were created with the crown area removed from all faces in order to eliminate possible biases associated with certain hairstyles evoking gender-biased information. Twenty-one participants (who did not participate in the main experiment) participated in the norming phase by rating the gender typicality of all faces on a 7-point scale (very masculine = 1, very feminine = 7). Presentation order of the faces was randomized for each participant. Six female faces were eliminated, as their ratings appeared to be less feminine (i.e. female faces that scored under 5 were eliminated; the remaining faces were scored appropriately: $M = 5.72$, $SD = .33$, range= 5.43 - 6.3). All 30 male faces were retained as they received clear ratings ($M = 1.58$, $SD = .26$, range: 1.23 – 2.47). These individual faces were then combined to make *male* and *mixed pairs of faces*. Female pairs of faces were not constructed for the experiment, as the interpretation of the presented masculine forms could not be grammatically interpreted as being female-specific (i.e., represented by female pairs of faces). More importantly, these female pairs of faces were avoided based on findings by Gygax and Gabriel (2008) who demonstrated that the presentation of both female and masculine forms in the same experiment directs readers towards a stronger male-specific representation of the masculine form. Female faces for mixed pairs were always presented on the left in order to avoid a male preferred response according to a possible left-side bias as illustrated in past studies using response scales in left-to-right languages (e.g., Gabriel et al., 2008). All pairs of faces were comprised of different faces.

In order to ensure that male and mixed pairs of faces were not processed differently due to some perceptual properties that we had not foreseen, we ran a pilot experiment. In this

pilot experiment, our experimental pictures were presented to another group of 27 participants who did not participate in the norming phase. Their task was to decide, on two blocks of trials, whether the presented pairs of faces were *of the same sex* in one block or *of different sex* in the other block, by indicating their responses with a *yes* or *no* button press. The block order was inversed for half of the participants. A repeated-measure ANOVA on correct response times (i.e. 94% of the data) showed no main effect of *block*, $F(1, 26) < 1$, *ns*, no main effect of *faces*, $F(1, 26) = 3.18^{13}$, *ns*, and no interaction, $F(1, 26) = 1.75$, *ns*, confirming the homogeneity of our experimental target stimuli in terms of perceptual properties.

4.3.2.3 L2 proficiency assessment

Participants' L2 proficiency levels were operationalized by a combined score of their own self-assessment and their performance scores on a given C-test. Commonly in a C-test, participants are given several distinct passages in which the second half of every other word is deleted except for the first and last sentences. The task is to restore the blanks in the allocated time. This procedure was developed as an effective measurement substituting cloze tests that were used in earlier years, and in recent years has been frequently used to measure language proficiency (Eckes & Grotjahn, 2006; Grotjahn, Klein-Braley, & Raatz, 2002). We employed the German C-test offered by onDaF (www.ondaf.de/) to test German proficiency. Score ratings on this test are considered equivalent to the Common European Framework of Reference for the levels A2 to C1. French proficiency was evaluated with Coleman's (1994) C-test. Four texts were chosen from each original version and 20 minutes were allocated to complete the task.

4.3.3 Design and procedure

The experimental task was conducted first in L1 and then in L2. Two experimental lists were created to ensure that a role noun would not appear in both languages for a given participant. The two lists were symmetrically different, in that if a role noun appeared in French in List 1, in List 2, it would appear in German. To avoid an imbalance of gender stereotypicality between languages, role nouns of similar strength of stereotype were always allocated to each language (see Table 1.1). Each list consisted of six female, six male and six neutral role nouns per language, resulting in 36 critical role nouns per list, with each role

¹³ If anything, participants were slightly faster (by 36 ms) to respond to mixed pairs of faces than to male pairs ($p = .08$).

noun appearing only in either language. Each role noun was presented four times per participant (cf. Gygax et al., 2012 and Oakhill et al., 2005 for a similar procedure): twice with a determiner (once followed by male pairs, once by mixed pairs of faces), and twice without. All experimental items were intended to elicit a *yes* response.

To trigger *no* responses, twenty filler role nouns that had a gender association by definition (e.g., *Großmütter* / *grand-mères* [grandmother]) were included. Half of the filler role nouns were male by definition, whereas the other half was female. These filler primes were also presented four times with their respective determiner allocations and face pairs. As these nouns were not ambiguous in terms of gender, including them prevented participants from responding *yes* throughout the experimental task without truly processing the role nouns and the target stimuli.

In each experimental trial, participants were first presented with a gender stereotypical role noun prime following a fixation point (1000 ms). The role noun was presented in the masculine plural form either in conjunction with a plural definite determiner (e.g., *die Ingenieure* / *les ingénieurs* [the engineers]) or without (e.g., *Ingenieure* / *ingénieurs* [engineers]). Participants were instructed to press the *yes* button after having read the presented role noun, which prompted the presentation of a picture of a pair of faces. Their task was to judge as quickly as possible with a *yes* / *no* button press whether the presented target face pairs could represent the prime role noun presented prior to the faces. Filler trials, which were randomized among experimental trials, followed the same procedure, and the role nouns within them were also presented either with or without a determiner.

The experiment finished with the completion of the C-test in each of the participants' respective L2. Following the C-test, participants were requested to assess their L2 competence in terms of their listening, reading, writing and speaking abilities in the L2 and to indicate the years and age of L2 acquisition by means of a self-administered questionnaire. Finally, participants were asked to translate the list of the 18 L2 role nouns they had seen for the main experimental task into their L1 in order to ensure they had properly processed the critical items.

4.3.4 Apparatus

The experiment was run on a Power Macintosh 4400 with the Psyscope software (Cohen, MacWhinney, Flatt, & Provost, 1993) connected to a button box to provide millisecond accuracy responses. Two buttons were labeled, one "*Ja*" (*yes*) and the other

“*Nein*” (no) for German-speaking participants and “*Oui*” (yes) and “*Non*” (no) for French-speaking participants. Items were presented on a computer screen and the “*Ja / Oui*” button was always pressed by the participant’s dominant hand. All participants were individually tested in a quiet room, with instructions being given in their respective native languages. They underwent a practice session in their L1 with four items in order to familiarize themselves with the task and procedure.

4.4 Results

Analyses were conducted on participants’ proportions of positive responses (i.e., *yes* responses) and their response times to the target pictures. Based on the results of the role noun translation task conducted after the experimental session, items in the L2 that were frequently unknown to each language group (fewer than 10% of the participants could provide a correct translation) were omitted from the analyses (*Schneider* [dress makers] and *Wahrsager* [fortune tellers] were removed from L2-French participants’ data and *diseurs de bonne aventure* [fortune tellers] from L2-German participants’ data).

Since C-tests are not necessarily comparable across languages, as pointed out in Sato et al. (2013), we analyzed the data separately for each native language group. To account for variations in language proficiency, we performed a median split on the C-test scores to create two language Proficiency groups (*low* vs. *high*). For the German-speaking native group, there were 23 low ($M = 84.65$, $SD = 5.33$) and 27 highly ($M = 93.89$, $SD = 2.28$) proficient participants, and for the French-speaking group, there were 26 low ($M = 54.77$, $SD = 10.14$) and 25 highly ($M = 81.04$, $SD = 10.30$) proficient participants. Mean differences in C-test scores between the low and high proficiency groups were significant for each native language group (German: $t(48) = 7.73$, $p < .001$; French: $t(49) = 9.18$, $p < .001$).

4.4.1 Proportion of Positive Responses

For the proportion of positive responses, we conducted both by-participants (F_1) and by-items (F_2) analyses. In the former (F_1), ANOVAs were conducted considering Task Language (German vs. French), Stereotype (female vs. male vs. neutral), Determiner (without determiner vs. with determiner) and Face Pairs (male vs. mixed pairs of faces) as within-subjects variables, and Proficiency (low vs. high) as a between-participants variable. In the latter (F_2) Proficiency, Determiner, and Face Pairs were treated as within-items variables and Task Language and Stereotype as between-items variables.

4.4.1.1 Native German group

Most importantly, there was a main effect of Face Pairs $F_1(1, 48) = 6.60, p < .05$; $F_2(1, 65) = 4.43, p < .05$; participants accepted male pairs of faces (.87, $SE = .03$) more often than mixed pairs (.79, $SE = .03$), thus confirming male biases found in previous research using verbal targets. There was also a main effect of Stereotype, $F_1(2, 96) = 18.25, p < .001$; $F_2(2, 65) = 5.98, p < .01$. Female stereotyped role nouns (.79, $SE = .03$) received fewer positive responses than male (.85, $SE = .02$) and neutral (.86, $SE = .03$) role nouns. Both main effects were qualified by a Stereotype by Face Pairs interaction, $F_1(1.334, 96) = 8.72, p < .01$; $F_2(2, 65) = 3.94, p < .05$. As illustrated by Figure 4.1, the face pair effect was stronger in the male stereotyped condition (.93 for male vs. .77 for mixed pairs of faces, $p < .001$) than in the other two conditions (female stereotype: .81 for male vs. .76 for mixed pairs of faces, *ns.*; neutral stereotype: .88 for male vs. .83 for mixed pairs of faces, *ns.*).

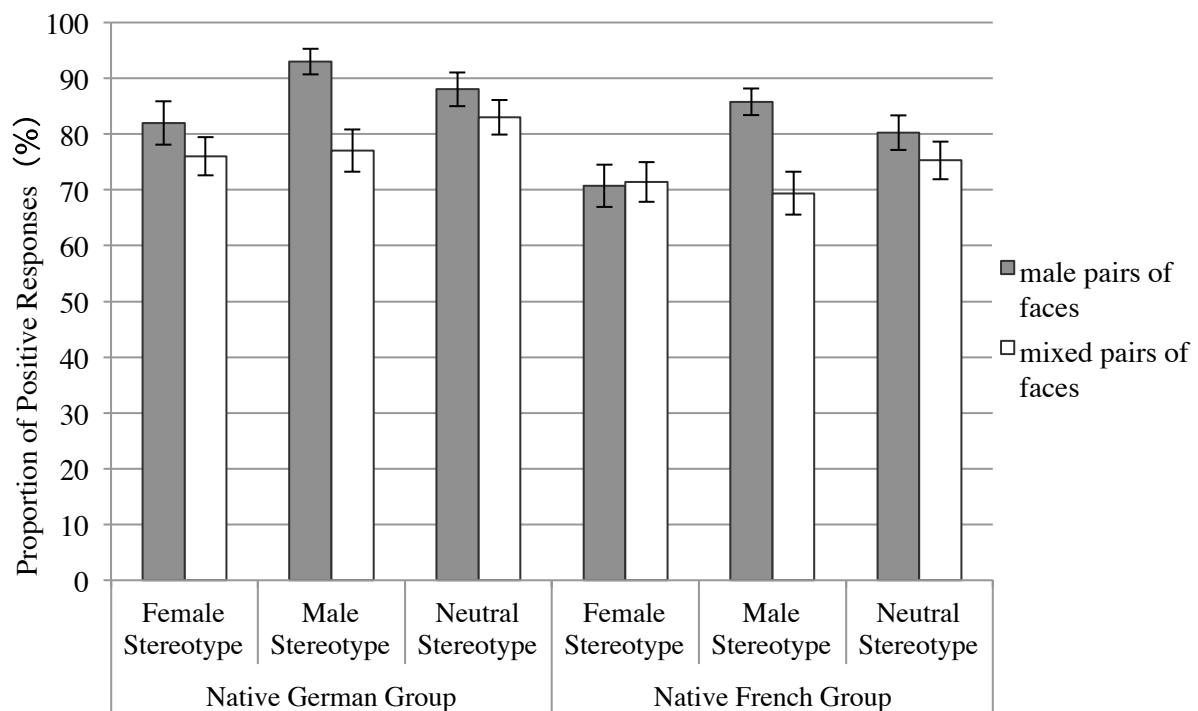


Figure 4.1 Proportion of positive responses to accept facial images for each native group in each stereotype condition (independent of task language). Error bars indicate standard errors.

Though there was no Determiner by Task Language interaction, which would have supported the notion of an effect of the German determiner *die*, there was a main effect of Determiner, $F_1(1, 48) = 5.25, p < .05$; $F_2(1, 65) = 5.87, p < .05$, participants responded

positively more often when role nouns were presented with (.85, $SE = .03$) than without (.81, $SE = .03$) the determiner.

Only one effect involved task language, namely a Face Pairs by Task Language interaction, $F_1(1, 48) = 10.96, p < .01$; $F_2(1, 65) = 9.67, p < .01$. Paired sample t-tests with Bonferroni correction showed that the Face Pairs effect was more pronounced in participants' L1 German, (male pairs .83, $SD = .21$; mixed pairs of faces .74, $SD = .30$; $p < .025$), than in their L2 French (male pairs: .83, $SD = .21$; mixed pairs: .78, $SD = .25$; *ns.*), suggesting that the male bias was stronger in their native language. No other main effects or interactions were significant (all $p > .05$).

4.4.1.2 Native French group

As was the case for the German sample, French participants showed a significant main effect of Face Pairs, $F_1(1, 49) = 11.22, p < .01$; $F_2(1, 64) = 29.97, p < .001$. Participants generally accepted male pairs of faces (.79, $SE = .03$) more often than mixed pairs of faces (.72, $SE = .03$), again confirming male biases found in the literature using verbal targets. There was also a significant Stereotype effect, $F_1(2, 98) = 15.49, p < .001$; $F_2(2, 64) = 4.72, p < .05$. Female stereotyped role nouns (.71, $SE = .03$) received fewer positive responses than male (.78, $SE = .03$) and neutral (.78, $SE = .03$) ones. As was for the German sample, both main effects were qualified by a Stereotype by Face Pairs interaction, $F_1(1.708, 98) = 16.49, p < .001$; $F_2(2, 64) = 18.68, p < .001$, suggesting that the main effect of Face Pairs was not equal across stereotype conditions. As illustrated by Figure 4.1, the face pair effect was stronger in the male stereotyped condition (.86 for male vs. .69 for mixed pairs of faces, $p < .16$) than in the neutral stereotyped condition (.80 for male vs. .75 for mixed pairs of faces), although absent in the female stereotyped condition (.70 for male vs. .71 for mixed pairs of faces).

There was an additional Proficiency by Face Pairs interaction effect, $F_1(1, 49) = 4.13, p < .05$; $F_2(2, 64) = 4.85, p < .05$. Low proficient participants showed a greater male bias reflected in a significantly greater proportion of positive responses for male (.80, $SD = .19$) than mixed pairs of faces (.69, $SD = .23$) ($p < .025$) as opposed to high proficient participants (male pairs of faces: .78, $SD = .23$; mixed pairs of faces: .75, $SD = .26$; *ns.*). Importantly, this effect surfaced independent of Task Language, which may indicate that the C-test in participants' L2 German could have been linked to their competence in their L1 French. A better competence in the L1 could be associated to an elevated awareness of the generic usage

of the masculine form, which could explain high proficient participants' ability to equally represent the faces.

4.4.2 Response Times

Only reaction times to positive responses (i.e., judgment times to accept the faces) were subject to analyses. Response times that were 2.5 standard deviations above or below the participant's mean (3.5%) were replaced by their cut-off values. To account for missing values (some participants gave only few positive responses) and to avoid the language-as-a-fixed-effect fallacy (Brysbaert, 2007; Clark, 1973), the response time data were examined by fitting linear mixed-effect models using the R software (R Core Team, 2013), with the *lmer* function in the *lme4* (Bates, Maechler, Bolker, & Walker, 2014) and *languageR* packages (Baayen, 2013). Contrasts for fixed effects were performed and corrected with Bonferroni corrections using the *glht* function from the *multcomp* package (Hothorn, Bretz, & Westfall, 2008). Degrees of freedom for *p*-values were corrected with Kenward-Rogers approximation using the *afex* package (Singmann, Bolker, & Westfall, 2015).

As suggested by Barr, Levy, Scheepers and Tily (2013) a model with a maximal random factor structure was adopted. Random intercepts and slopes were varied for participants and items in order to account for the variance in performance created by the factors Baayen, Davidson and Bates (2008). Random slopes were eliminated if their removal did not result in a significant amelioration of the model or if they did not converge. For the native German group, the final model included random slopes for Task Language at both participant and item level. For the native French group, the random structure included random slopes for Face Pairs at participant level. Similar to the ANOVA analyses for the response proportions, the predictors entered as fixed effects in the reaction time model included Face Pairs (male vs. mixed pairs of faces), Stereotype (female vs. male vs. neutral), Task Language (French vs. German), Determiner (with determiner vs. without determiner) and Proficiency (high vs. low). All predictors were sum coded (1, -1) and were entered in a stepwise manner according to an initial null model including only random factors. To determine the inclusion of the predictors in the model, models were compared using *log-likelihood ratio* tests. A more complex model including the particular predictor was compared to a simpler model without the predictor in it. If an integration of a predictor significantly improved the model, it was retained as a predictor within the model. For both language groups, fixed effect correlations for the best fitting models were less than .5, and variation inflation factors were less than 1.4

indicating that collinearity was not a problem. We present the final models for the native German group and the native French group respectively.

4.4.2.1 Native German group

The final model included main effects of Face Pairs, Stereotype, Proficiency, Task Language and Determiner as well as a Face Pairs by Stereotype interaction (see Table 4.1).

Table 4.1 Fixed effects of the reaction time model of the German group with participants and items as random factors, and by-participant and by-item random slopes for task language.

	β	SE	t -value	p -value	
(Intercept)	959.66	67.53	14.21	< .0001	***
Face Pairs (men)	-97.60	7.92	-12.33	< .0001	***
Stereotype (female)	46.38	15.74	2.95	< .001	**
Stereotype (male)	1.76	15.61	.11	0.96	
Task Language (German)	82.19	19.79	4.15	0.17	
Determiner (with)	-16.24	7.78	-2.09	0.99	
Face (men) by Stereotype (female)	15.31	11.14	1.38	< 0.001	**
Face (men) by Stereotype (male)	.42.36	10.97	-3.86	< .0001	***

* $p < .01$ ** $p < .001$. *** $p < .0001$

The main effect of Face Pairs was significant $F(1, 5865.54) = 151.98, p < .001$, confirming the male bias in past studies. Male pairs (862 ms) were agreed with faster than mixed pairs (1057 ms) of faces. There was also a significant Stereotype effect $F(2, 32.2) = 5.72, p < .01$. Pairwise comparisons showed that face pairs following neutral stereotyped role nouns ($M = 911$ ms) were responded to significantly faster than when following female stereotyped role nouns ($M = 1006$ ms, $p \leq .001$). No comparison involving male stereotyped role nouns ($M = 961$ ms) was significant. The main effect of Task Language was also significant, showing that response times were slower in participants' L1 German ($M = 1042$ ms) than in their L2 French ($M = 877$ ms). These results could suggest a general sequence effect where participants were more familiarized with the task having started in their L1 and switching to their L2 for the second half of the experiment. A main effect of Determiner $F(1, 5841.93) = 4.35, p < .05$ suggested that responses were faster when the primes were presented in conjunction with a determiner (943 ms) than without (976 ms). This effect indicating that

role nouns without determiners are more difficult to process, will further be discussed in the Discussion section. Crucial for the study, a Stereotype by Face Pairs interaction surfaced, $F(2, 5855) = 7.7$, $p < .001$. Across all stereotype conditions, participants accepted male pairs of faces faster than mixed pairs of faces. In particular, the difference was more pronounced for male stereotyped role nouns (219 ms) than for neutral (131 ms) or female (163 ms) stereotyped role nouns. Acceptance to mixed pairs of faces was facilitated the most for neutral stereotyped role nouns. Importantly, no interaction term including Task Language or Determiner improved the final model.

4.4.2.2 Native French group

The final model included the main effects of Face Pairs, Stereotype, Proficiency, Task Language as well as the interactions Face Pairs by Stereotype and Proficiency by Task Language (see Table 4.2).

Table 4.2 Fixed effects of the reaction time model of the German group with participants and items as random factors, and by-participant and by-item random slopes for task language.

	β	SE	t -value	p -value	
(Intercept)	1051.40	54.44	19.313	< .0001	***
Face Pairs (men)	-87.54	17.34	-5.048	< .0001	***
Stereotype (female)	42.10	14.6629	2.871	< .01	**
Stereotype (male)	0.71	14.47	0.049	0.96	
Task Language (German)	10.40	7.50	1.386	0.17	
Proficiency (High)	0.58	46.82	0.012	0.99	
Face (men) by Stereotype (female)	23.94	9.26	2.585	< 0.01	**
Face (men) by Stereotype (male)	-50.99	8.98	-5.678	< .0001	***
Task language (French) by Proficiency (High)	42.93	7.53	5.704	< .0001	***

* $p < .01$ ** $p < .001$. *** $p < .0001$

As was the case for German speakers, the main effect of Face Pairs revealed that participants were faster to accept male pairs of faces (963 ms) than mixed pairs of faces (1139 ms), in line, again, with a strong male bias, $F(1, 98.4) = 25.46$, $p < .001$. A significant Stereotype effect suggested different levels of difficulty in representing each stereotype, $F(2, 32.9) = 5.66$, $p < .01$. Similar to the native German group, pairwise comparisons revealed that

face pairs following neutral stereotyped role nouns ($M = 1008$ ms) were responded faster than when following female stereotyped role nouns ($M = 1094$ ms, $p < .01$). No comparison involving male stereotyped role nouns ($M = 1052$ ms) was significant. This effect was further qualified by a significant Stereotype by Face Pairs interaction, $F(2, 8549.87) = 16.22$, $p < .001$. Similar to the native German group, the male pairs of faces were responded to faster than mixed pairs of faces for all stereotypes. However, the difference was again more pronounced for male stereotyped role nouns (223 ms) than for neutral (107 ms) or female (132 ms) stereotyped role nouns. Again, the acceptance to mixed pairs of faces was most facilitated when the faces followed neutral stereotyped role nouns.

Although neither the effects of Proficiency, $F(1, 98.81) = 0$, *ns.*, nor Task Language, $F(1, 8668.38) = 1.92$, *ns.*, reached significance, their interaction was significant $F(1, 8597.72) = 32.49$, $p < .001$. Namely, high L2 proficient participants were likely to respond (agree) faster in their L1 French than in their L2 German. In contrast, low L2 proficient participants took longer time in their L1 French than they did in their L2 German. Such results could suggest that low participants were more likely to speed up in the second half of the experiment (conducted in the L2) due to familiarity of the experimental procedures even when the task language (L2) could have presented them with a more complex task. In contrast, high proficient participants may have been more willing to engage in the task in their L2, and may have taken longer time to finish the second half of the experiment regardless of the familiarity with the experimental procedures.

4.5 Discussion

The aim of the present study was twofold. First, we aimed to provide more compelling evidence of the interaction of stereotypical and grammatical gender information during lexical processing of role noun cues. The impact of gender stereotype information has often been assumingly overshadowed by grammatical gender information in past studies, resulting in some uncertainty as to how stereotype information actually influences the interpretation of the masculine form. While most experimental tasks in past studies relied on verbal targets, we argue that this may have reinforced the grammatical and morphological properties of the role noun cue being tested, resulting in strong, yet less generalizable grammatical-based representations. We therefore proposed a new experimental approach using visual targets to gauge these effects. The second aim was to evaluate how linguistic encoding of gender concepts in different languages shape and shift gender representations. Employing French and

German bilinguals, gender interpretation differences between the two languages that were potentially prompted by differing surface forms were tested to assess this issue. More specifically, we further explored the rather inconclusive effects of a female bias associated to the German determiner *die* (gender neutral in the plural, but same surface form as the feminine singular determiner). Given the possible different gender biases present in the determiners of the two languages, as well as associated stereotypical and grammatical gender information that are activated during representation processes, we expected shifts in representations between participants L1 and L2, thus providing for the first time direct evidence as to the prominence of linguistic cues in shaping gender representations.

Results for both language groups revealed a consistent main effect of Face Pairs, where the representation of male pairs of faces was facilitated over mixed pairs of faces. This preference to accept the male face pairs reflects the general ease in interpreting role nouns marked for the masculine form as being male-specific rather than generic. Although the surface form of the masculine grammar can theoretically be detached from its semantic association *masculine* = *men*, it nonetheless boosted the activation of semantic properties associated to the male gender. This was true even though participants were presented with visual targets. Importantly, this male bias was persistent even when considering that our pilot experiment on the facial images showed a slightly faster, although not statistically significant ($p = .08$), tendency to process mixed pairs of faces. Our results therefore suggest that a strong male bias is indeed generated by the grammatical masculine form, and is not simply an artifact of the experimental task employed in previous studies.

For both language groups, participants' responses to facial targets were influenced by the stereotypicality of the role nouns, with neutral stereotyped role nouns generating processing facilitation of following facial targets. In contrast, facial targets following role nouns with a specific gender stereotype (i.e., male and female) were more difficult to process, as readers may have needed to construct more precise representations that incorporated gender information. In particular, processing facial targets after having processed female stereotyped role nouns proved to be most difficult, both in terms of reduced positive response proportions and inflated response times. We believe this to be indicative of interference between the grammatically masculine form and the role noun's female stereotypicality. Namely, both sources of information compete, increasing processing time. In contrast, greater positive response proportions for targets following male stereotyped role nouns suggest that

the congruency between the grammatically masculine gender and stereotypical gender facilitated participants' construction of mental representations.

This effect was further qualified by a consistent Stereotype by Face Pairs interaction for both the German and French group. This interaction indicated that participants' acceptance to face pairs changed in function to the stereotypicality of the role noun preceding it. Male stereotyped role nouns triggered the greatest facilitation to accept male pairs of faces, whereas neutral stereotyped role nouns triggered the greatest facilitation to accept mixed pairs of faces. These results support the idea that when reading a gender associated role noun such as *nurses* (*Krankenpfleger*_{German}, *infirmiers*_{French}), or *bosses* (*Arbeitgeber*_{German}, *patrons*_{French}) in a grammatical gender language, gender stereotypical information is immediately activated as part of the information associated with the role noun. As we did not embed our primes within sentences, our results suggest that this activation is made at the lexical access, with discursive text elements not needed to guide the activation of gender stereotypical information. Although we did find evidence that the masculine form was highly influential in guiding the representation toward a male-dominant representation as found in previous studies, we also documented that readers rely on immediate stereotypical information.

However, these results do not necessarily speak to whether, and to what extent, grammatical gender or stereotypical information has a greater influence over gender representation, as discussed in some discourse-based studies (Irmen, 2007). They mainly support the idea that both are activated at an early stage (i.e. lexical access), a claim that contrasts those of anaphor resolution studies that suggest an activation at later stages of comprehension (e.g., Esaulova et al., 2013; Irmen, 2007). The absence or weak indications of strong and immediate stereotype effects in past studies could be attributed to several reasons. First of all, past research has frequently relied on verbal primes *and* verbal targets (e.g., Gygax et al., 2012; Gygax & Gabriel, 2008) to substantiate a persistent effect of the masculine form as specifically referring to men, with the effects of stereotype being only modest. The present study however, demonstrated that the apparent lack of stereotype effects could be attributed to the tasks used to investigate these issues. We believe that by using facial images as targets, we went beyond simple language-on-language task effects. Essentially the conceptual nature of stereotypes may have made them better candidates for non-verbal tasks which made it possible to delineate the true and noteworthy interaction between grammar and stereotypes when constructing a representation of gender. Another plausible argument for the absence of stereotype effects in past studies can be accredited to

the nature of stereotype information, which dwindles rapidly as readers process discourse. Consequently, its effects did not clearly surface in previous studies on text comprehension. In the present study, the lexical-based paradigm may have allowed stereotype effects to surface before fading away, as they would have in a discursive context. Such a view may also support the reason for grammatical gender information to show a greater impact in most studies on the topic.

In terms of the impact of language on shaping gender representations, the two language groups showed similar representation regularities in both their L1 and L2. This was rather unexpected given that we had anticipated the male bias to be reduced when participants processed the role nouns in German, due to its female-equivalent determiner. In fact, the German determiner did not elicit any substantial effects. Although there was a modest trend for mixed pairs of faces to be accepted more often when following female and neutral stereotyped role nouns (proportion of positive responses) when adding the determiner *die* for native German readers in L1, it did not lead to statistically significant effects. Nonetheless, these results are in line with many gender representation studies suggesting that the male bias exerted in grammatical gender languages through the use of the masculine form is strong and appears to govern the process of comprehension. Although we cannot definitively refute the phenomenon, the male-attenuating effect in German documented by Rothermund (1998) appears to be superficial in relation to the male-bias exerted by masculine forms, though interpretable as generic. The fact that Garnham et al. (2012) found an effect of *sie* [they_{Female}], was most likely due to the fact that they combined *die* and *sie*, both feminine equivalent, which offered a cumulative effect in deterring readers' attention from the role nouns' masculine form.

Nevertheless, our results did show a main effect of Determiner for our German group, whereby role nouns with a determiner facilitated responses to targets. These effects could be explained by the different rules associated to German. For instance, in French although a noun must always be accompanied by a determiner even when a general statement is being made (e.g., *Les infirmiers* doivent soigner les personnes. [*Nurses* need to care for people.]), in German, a noun can be presented both with and without a determiner (*Krankenpfleger* müssen sich um Menschen kümmern. [*Nurses* need to care for people.] vs. *Die Krankenpfleger* müssen sich um Menschen kümmern. [The nurses needed to care for people.]) which denote different meanings. The presence of *die* more clearly specifies that the role noun refers to a *group of people*, and not to the general activity represented by the role

noun, consequently facilitating subsequent associated targets. In this regard, our German group may have constructed different representations according to whether the role noun was presented with or without a determiner.

Although the male-attenuating effect of the German determiner was not observed, we nonetheless observed a Task Language by Face Pairs interaction surfacing in our German group's responses. This effect suggested that the male bias was in fact more persistent in participants' L1 German than in their L2 French. This is crucial given that their dominant language exerted a greater male bias than their less fluent L2, despite having a better understanding and command of the language and the different interpretations of the masculine form in their L1. Interestingly, for the French group, a Proficiency by Face Pairs interaction surfaced on response proportions. These results gave rise to the possibility that higher proficiency as measured by L2 C-tests also gauged linguistic competence in one's L1. If this were true, the results suggest that competence in one's L1 is linked to better command in using the generic usage of the masculine form. Together, these results suggested that the male bias appears to stem from L1 for grammatical gender language readers. When switching languages, reader's L2 proficiency provides a meaningful indicator as to their command of the generic usage of the masculine form in both of their language and the extent it surfaces on their representations. Such an account is in line with bilingual processing theories that suggest that the languages of a bilingual are non-selectively activated even when only one language is being used for language comprehension processes (e.g., de Groot, Delmaar, & Lupker, 2000; Dijkstra & van Heuven, 1998). Effective processing mechanisms will then require a suppression of the unused language.

The findings of our study are critical as they illustrate how language can fundamentally guide and influence how information is represented. In our study, we observed a heavy reliance on the grammatical masculine form, contributing in shaping male-dominant representations across (more or less) all stereotypes, which was at odds with the idea that the masculine is the *unmarked* gender in grammatical gender languages. These results lend support to the idea that grammatical markings may well direct (or bias) our attention to particular categories. The masculine form most likely makes the male concept more accessible to readers. Although our cognition of gender itself may not be fully influenced by grammatical gender, and this is an empirical question, our social cognition may well be, given that the concept of gender, especially that of male, is enhanced in grammatical gender language readers. These tendencies may then result in shifting or influencing our social

perceptions of gender-stereotyped occupations, guiding readers to integrate a representation that is advantageous for men (Braun et al., 2005; Irmen & Köhncke, 1996).

4.6 Conclusion

Using a linguistic-visual paradigm, the present study showed that readers automatically activate gender-associated information when reading gender stereotypical human referent role nouns. The activation of such information immediately takes place at a lexical level when readers encounter a role noun. Though morphological markings such as the default masculine form in French and German appear to be central when constructing mental representations of gender, stereotype information also plays an interactive role in influencing readers' mental representations. The latter interaction is particularly apparent in the cumulative effects of stereotype and grammar when readers encounter male stereotyped role nouns. Future studies may want to further examine the possibilities of suppressing such male-dominant properties, though they appear to be relatively robust.

Chapter 5 : Study III

Women do not operate on patients, men do: Activation of gender information during subtle gender occupational descriptions

5.1 Abstract

Study II revealed that in grammatical gender languages, stereotype and grammatical information were immediately activated when readers read role nouns. While both stereotype and grammatical information showed an intricate interaction, grammatical gender was found to ultimately influence representation, facilitating the representation of the male gender from the grammatical gender cue (interpreted “specifically”). These findings suggested that role nouns automatically trigger gender representations as these gender features are incorporated as part of their lexical representations. Following these findings, the present study sought to assess the robustness in which these gender inferences are generated by comparing cues that provide strong and weak gender encoding devices during text comprehension.

In the present study, we report two eye-tracking experiments in L1 French and L2 English that tested whether these gender inferences could be activated when the source of gender information did not designate a specific gender category (as in a role noun), in a match-mismatch paradigm using eye-tracking measures. Experiment 1 replicated past accounts of an immediate activation of gender through gender-biased role nouns, where information about the stereotypical gender and grammatical gender were both activated. In Experiment 2, occupational descriptions (e.g., *operating on the patient*) associated to a certain gender replaced these role nouns. Results suggested that occupational descriptions are not strong enough cues to spontaneously activate gender information. Instead, gender information activation was only elicited when readers encountered pronouns that triggered an antecedent search suggestive of a backward inferencing mechanism.

5.2 Introduction

Reading comprehension consists of constructing mental models that link explicit textual information with implicit information readers infer from their world knowledge (Garnham & Oakhill, 1996; Johnson-Laird, 1983). However the mechanisms and conditions for which different types of inferences are generated vary, as some have been found to be triggered automatically¹⁴ (e.g., Oakhill, Garnham, & Reynolds, 2005), while others are dependent on contextual (e.g., Duffy & Keir, 2004) and grammatical (Gygax, Gabriel, Sarrasin, Oakhill, & Garnham, 2008; Irmen & Roßberg, 2004; Irmen, 2007) influences. This study focuses on how gender cues with different magnitudes of saliency instigate inferences for gender representation.

Evidence is accruing from recent research to suggest that readers immediately make use of conceptual gender information when reading a gender-associated role noun (Banaji & Hardin, 1996; Cacciari, Corradini, Padovani, & Carreiras, 2011; Oakhill et al., 2005). While conceptual knowledge refers to “a person’s cognitive representation of a category” which is essentially used during processes associated to reasoning, problem solving and language use (Barsalou, 1993), in terms of gender, it is also linked to knowledge of basic principles as to whether certain roles are taken on by men and women. Specifically in the context of the following experiments, we refer to gender concepts as being *stereotypical* which reflects readers’ social expectations about how women or men should and are expected to take on these social roles (e.g., the assumption that women are better caregivers than men, men make better leaders than women, etc.).

The effects for the activation of conceptual gender information can be observed during the representation of biologically definitional (e.g., female: *sister*, male: *father*) and stereotypically biased role nouns (e.g., female stereotype: *dressmaker*, male stereotype: *surgeon*, from Gabriel, Gygax, Sarrasin, Garnham, & Oakhill, 2008). In the former definitional role nouns, gender is semantically determined whereas in the latter stereotypical role noun, readers are not *necessarily* confined to allocating a specific gender. Nonetheless, general stereotypical assumptions (e.g., surgeons are more likely to be men) have been found to lead readers to designate a specific gender for stereotypical role nouns, which becomes

¹⁴ As previously noted, the term automaticity has been extensively debated as to the notions it implies (e.g., Bargh, 1999). In this study, the term *automatic* is used to denote the immediacy in which activation occurs as well as to illustrate the fact that it occurs without the reader’s intent.

incorporated as part of their mental representation (e.g., Carreiras, Garnham, Oakhill, & Cain, 1996; Oakhill, et al., 2005; Reynolds, Garnham, & Oakhill, 2006).

Studies using word-based paradigms have demonstrated that stereotypical gender information establishes strong priming effects between a gender stereotypical role noun prime (e.g., *surgeon* / *nurse*) and a gender definitional target word (e.g., pronouns: *she* / *he*, Banaji & Hardin, 1996; kinship terms: *mother* / *uncle*, Oakhill et al., 2005). In their studies, when there was a gender match between the stereotypicality of the prime and gender designation of the target words, reaction times were shown to be faster than when there was a mismatch between the primed and targeted gender. Following these results showing gender priming effects, research in sentence comprehension has shown that the activation of gender stereotypical assumptions may lead to reading difficulty.

For example, Kennison and Trofe (2003) investigated how gender stereotypes associated to role nouns would influence pronoun resolution during sentence comprehension. In a self-paced moving window paradigm, participants read sentence pairs including a gender biased role noun in the first sentence (e.g., male stereotype: *executive*, female stereotype: *secretary*), followed by a gender matching or mismatching pronoun (*she* / *he*). Reading times were inflated if the gender of the role noun stereotype and pronoun mismatched, indicating a processing difficulty. Similarly, Gygax et al. (English data, 2008) also reported similar findings that indicated that readers' plausibility for sentence continuations including a gender reference (e.g., *some of the women* vs. *some of the men*) was greater if there was a gender congruency with an aforementioned gender stereotypical role noun (female stereotype: *nurses*; male stereotype: *surgeons*; neutral stereotype: *pedestrians*). The gender mismatch effects demonstrated in these studies highlighted that when the information specifying the gender such as a pronoun did not match the readers' assumptions about the probabilistic gender of the role noun reference, comprehension difficulty arose. The difficulty was considered to be stemming from readers' efforts to compensate for the gender violation if the textual specification of gender contradicted their pre-activated assumptions about a probabilistic gender of the role noun. Such a view assumes that a specific gender was already represented before readers read the information specifying the actual role noun gender (i.e., pronoun). In this regard, the locus of gender activation was situated precisely at the role noun cue, and was not only triggered when readers encountered the anaphor that required an antecedent search.

Based on these lines of evidence, some researchers have argued that stereotypical role nouns may yield elaborative inferences that are made in a forward direction (e.g., Oakhill et

al., 2005; Pyykkönen, Hyönä, & van Gompel, 2010). Namely, this type of inference is generally elicited to enhance the understanding of the text by enriching the mental model, although they may not necessarily be needed for the purpose of achieving text coherence (Garnham, 2001; Johnson-Laird, 1983). Garnham, Oakhill and Reynolds (2002) argue that among other information, these inferences about the gender stereotypical characteristics of a person reference are contained within the representation of the role noun itself, which then automatically activate gender information when readers encounter the role noun. Readers will update any contradicting information if need be. Such an explanation also fits well with the lexical reinterpretation model developed by Hess, Foss and Carroll (1995) who proposed that a single word allows all associated concepts (or at least those that are typical) to be accessible to the reader. Interpretations are constantly updated into the representation as new information is encountered throughout the text. For instance reading a word like *English major* in a text would essentially activate related aspects such as “poetic” or “adult”. If however contextual information portrays someone that does not necessarily fit the assumed characteristics (e.g., a student in a technical course on computing) readers will effectively modify their representation to concur with the local information in the text.

Contrary to English, languages such as French illustrate a different picture of gender representation, as the languages themselves consist of gender as a grammatical feature. In these languages, gender is grammatically marked as a morphological feature that explicitly denotes the categorical sex of the referent. However, the allocation of the sexes onto the grammatical marking does not occur on a one-to-one basis, which has been found to create processing or interpretation challenges. Importantly, complications may also arise given that this grammatical information may interact with readers’ held gender stereotypes. For example, Vigliocco and Franck (2001) investigated the interaction of grammatical and conceptual gender information employing grammatically gender marked words that can refer to both sexes (i.e., epicenes¹⁵). The difficulty in understanding epicenes is based on the fact that the grammatical gender marking is not directly related to its conceptual gender (e.g., *la victime*_{Feminine} [the victim – for both male and female]). Participants in their study were presented with passages that introduced an epicene that followed a prior specification of its conceptual gender by a person name (e.g., *Fabien* for a man vs. *Fabienne* for a woman). When instructed to produce a gender marked adjective (e.g., *sourde*_{Feminine} vs. *sourd*_{Masculine} [deaf]) where the inflection needed to correspond to the epicene, participants were found to

¹⁵ These gender marked words that can refer to both sexes have also been referred to as bi-gender nouns in past studies (Cacciari et al., 2011; Cacciari & Padovani, 2007).

make more inflection errors when the grammatical gender of the epicene mismatched the denoted conceptual gender information (i.e., errors such as *sourd*_{Masculine} for *la victime*_{Feminine} when prior texts mentioned *Fabien*, a man). Although their study did not explicitly touch on stereotypical gender, their results indicated that gender conceptual information had a facilitative impact on accuracy when there was a match with grammatical information, whereas it hindered processes where a mismatch was found.

Crucial to the present study is the notion of the *generic* interpretation. In most grammatical gender languages such as French, while the feminine form unequivocally refers to the female sex (e.g., *chirurgienne*_{Feminine} [female surgeon]), the masculine form (e.g., *chirurgien*_{Masculine} [surgeon]) has multiple denotations, referring to (a) the male sex, (b) both sexes or (c) used in cases where the sex is unknown or unimportant. Studies have shown that the multiple interpretation possibilities for the masculine form may pose difficulty in the gender representation process. In the aforementioned study by Gyga et al. (2008), the same experimental manipulation was applied in the French and German languages where the role nouns were presented in the masculine form that could have both male-specific and generic interpretations (e.g., French: *chirurgiens*_{Masculine}, German: *Chirurgen*_{Masculine} [surgeons]). Contrary to the stereotype-consistent effects found in their English data, the authors reported that readers found references to men (i.e., some of the men) to be more plausible following the stereotypical role nouns irrespective of the gender stereotypicality of the role nouns. The authors argued that during the interpretation of the masculine form, grammatical gender information governs representations processes that overrule simultaneously activated stereotype information.

The aim of the present research was to further complement past research indicating that reading a gender stereotypically biased role noun activates a specific gender within the mental representation. In cases where only stereotypical information is available, readers link role nouns to a probabilistic gender as it consists part of the representation of the lexical item. This activated gender is then carried as part of the representation during the course of reading. However, when gender information is complemented with grammatical gender information, as is in grammatical gender languages, linguistic parameters impose constraints on the specification of gender. In both cases, gender information related to the role noun becomes a salient dimension during the construction of the mental model.

These assumptions provide explanations to ascertain that role nouns make salient cues that initiate an automatic activation of gender. Naturally, a question that arises from such assumptions is whether these gender inferences could still be generated even if the conditions

do not present a heightened cue to a categorical gender. In other words, if these specific lexical items (i.e., role nouns) that elicit strong gender inferences are omitted, yet the presence of a person reference is still made clear, it seems probable that readers would still activate and encode gender into their representation despite its reduced effects. A reference to a person for instance can still be made possible without *explicitly* having to mention them (i.e., role noun) by illustrating environmental settings that denote the presence of people in general (e.g., The train was crowded) or for more detailed purposes, illustrating a description of the activity which the person is engaged in (e.g., walking in a crowded city). In terms of gender biases, it is plausible to assume that general activities are related to a specific gender (e.g., female stereotype: *knit a sweater*, male stereotype: *build a house*), as is the case for role nouns. However, given that descriptions may not require readers to activate a certain categorical label, they are most likely less salient in nature in terms of representational purposes.

This view that less prominent markers may serve as gender activating devices is related to Andersen and Klatzky's studies comparing stereotypes and stereotypical traits (e.g., being *conservative* and *conscientious*; Andersen, Klatzky, & Murray, 1990; Andersen & Klatzky, 1987). The authors defined stereotypes as being stronger encoding devices that carry whole feature information about a person or social groups which contrasts to stereotypical activities that carry only particular dimensions of person concepts, consequently offering weaker cues than categorical labels. In the present study, we aim at approaching gender information cues under a similar rationale by comparing the magnitude of stereotypical role nouns that provide strong encoding devices for gender information, with those of occupational descriptions that are characteristic of role nouns but encode gender to a lesser extent than the former. Occupational descriptions may serve as suitable cues for weak gender encoding devices given that the origins of gender stereotypes are viewed as being rooted in the division of labor and social roles between the two sexes (Eagly, 1987; Eagly & Wood, 1999). It would be reasonable to assume that people may perceive occupational descriptions as being associated to a specific gender. Essentially, investigation of these processes allows for the examination of how non-categorical information may contribute to the generation of specific inferences.

Based on a similar line of reasoning, Reali, Esaulova and Von Stockhausen (2015) presented German sentences including stereotypical descriptions (e.g., *M. F. repairs and produces pieces of furniture*) followed by an anaphoric pronoun (*he / she*) in an eye-tracking setup. They found mismatch effects on fixation times and regressions when the anaphor gender violated the stereotypicality of the description gender. These effects were found at

early, middle and late stages of processing suggesting effects of stereotype information. These results were interpreted as showing gender reference activation when readers read descriptions of stereotypical professional roles.

Importantly however, these studies have not sufficiently highlighted the extent to which language will influence the representation process. With past research demonstrating the different representations bilingual readers construct based on their two languages, it is plausible to speculate that the impact of these cues varying in gender saliency should depend on the language in which these cues are expressed. This is because for French readers, monitoring gender information is crucial, as rules governing grammatical gender in French requires readers to attend to gender information cues during language processing. Such requisites contrast to comprehension in English, where no such linguistic obligations are found. In this regard, examining whether the impact of cues varying in magnitude of gender saliency differ among French readers when reading in French and in English may provide a better insight into how these gender cues function during representation processes. Representations may ultimately differ when readers comprehend in French that imposes close monitoring of gender information and in English that does not impose such attention allocation. Past research investigating the gender representation processes of French (L1 [first language]) – English (L2 [second language]) bilinguals have shown that the effects of the masculine form in the L1 French (as seen in Gygax et al., 2008) elicited a strong male bias to be carried over to the bilinguals' L2 English when participants' L2 proficiency was found to be low, lacking linguistic control (Sato, Gygax, & Gabriel, 2013).

To this end, the aims of the following experiments were twofold: (1) To examine the magnitude of stereotypical role nouns as a gender cue in French-English bilinguals, and (2) To investigate whether the magnitude of these role nouns would differ in contrast to cues with weaker encoding strengths (i.e., gender occupational descriptions) that do not serve as a categorical label. In what follows, two eye-tracking experiments tested anaphor resolution processes during the comprehension of short passages in L1 French and L2 English that presented gender stereotypical information followed by a gender-specifying pronoun (i.e., *elle* - *she* / *il* - *he*). In Experiment 1A, gender stereotypical information was conveyed through a presentation of a gender stereotypical occupational description (e.g., female stereotype: *shorten the trousers*; male stereotype: *operate on the patient*), followed by a corresponding gender emphasizing role noun (e.g., female stereotype: *dressmaker*, male stereotype: *surgeon*). On the other hand, in Experiment 1B, only occupational descriptions were presented as a source of gender information without further reinforcement of gender

stereotypicality by an associated role noun. In both experiments, the indication of gender information was expected to yield gender inferences elaborately when role nouns or occupational descriptions were read. Following past research, this immediate activation of gender was expected to elicit a match-mismatch effect on eye-fixations when encountering a gender-specifying pronoun (i.e., *elle* / *she* and *il* / *he*).

5.3 Experiment 1A

In this experiment, the activation of gender information was assessed in L1 French and L2 English. Participants were presented with stereotypically biased occupational descriptions and their corresponding role nouns that were followed by a pronoun specifying the gender of the person reference (e.g., *shorten the trousers – dressmaker – he / she*).

In L1 French, the role nouns were presented in the masculine singular form that could have both a male-specific and generic interpretation. Although the generic interpretation in the singular forms have never been empirically tested, past research has supported the notion that readers are more likely to interpret masculine forms as being male-specific (e.g., Gygax et al., 2008). Following such results, we expected that the masculine feature of the role nouns would trigger a male-specific interpretation, with participants exhibiting a mismatch effect if the following pronoun information specified a female gender (i.e., *elle* [she]). A certain level of interference was also expected to surface if the stereotypical and grammatical information of the role noun clashed (i.e., masculine grammar with female stereotype in cases such as *couturier*_{Masculine} [dressmaker]). Alternatively, in L2 English, a general gender-mismatch effect was expected to surface if the gender stereotype conveyed prior to the pronoun violated the specification of gender by the pronoun (e.g., *dressmaker - he*; *surgeon - she*). However as suggested by Sato et al. (2013), a greater male-bias was expected to take effect for participants with low levels of L2 English proficiency (as assessed by an L2 evaluating C-test) as they would show less control over the processing language.

5.3.1 Method

5.3.1.1 Participants

Twenty-eight L1 French-L2 English speakers (21 women, seven men)¹⁶ from the University of Fribourg, Switzerland participated in the study for course credits.

¹⁶ As past research has not found substantial effects of participant sex on gender representation, this variable was not controlled for in the two experiments.

5.3.1.2 Materials

5.3.1.2.1 Experimental stimuli

Thirty-six experimental passages were prepared for the experiment (see French [a] - [c] and English [a]-[c] below). Each set of passages comprised of three parts. The first part introduced a normed gender stereotypical occupational description, which was followed by the second part that mentioned a female (e.g., *dressmaker*), male (e.g., *surgeon*) or neutral (e.g., *pedestrian*) stereotypical role noun (all role nouns were employed from Gygax et al., 2008 that had been normed for gender stereotypicality in Gabriel et al., 2008). The stereotype information was expected to be encompassed within the lexical representation of the role noun, which would be activated at the point role nouns are encountered in the text. Gender stereotypicality between the role nouns and the preceding occupational description matched at all times, reinforcing the conveyed gender stereotypicality. Finally in the third part, either a female (i.e., *elle – she*) or a male (i.e., *il – he*) pronoun that referred back to a denoted person reference was presented.

The pronoun was immediately followed by a modal verb (i.e., a verb that is used to express sentence modality as in *she believed, he hoped*). Across the 36 experimental passages, 12 were of female, 12 were of male and 12 were of neutral stereotype. All passages had a French and an English version that were translated from one another. Below, we present an example of an experimental item with regions of interest that will be analyzed in the analyses indicated in italics (further described in the analyses section).

OCCUPATIONAL DESCRIPTION REGION:

French (a) *En raccourcissant les pantalons,*

English (a) *While shortening the trousers,*

ROLE NOUN REGION:

French (b) le *couturier* paraissait préoccupé.

English (b) the *dressmaker* seemed confident.

PRONOUN REGION:

French (c) *Elle / Il* savait que la tenue conviendrait

English (c) *She / He* knew that the outfit would fit.

All occupational descriptions used in the study were normed prior to the actual experiment to ensure that they were typical examples of their corresponding role nouns. They

all portrayed stereotypical actions of their corresponding role nouns, (e.g., *Shortening the trousers* was the occupational description for *dressmaker*). All descriptions were constructed using the same structure, beginning with a main verb in the present participle preceded by the conjunction “*En / While*”. This sentence structure allowed the presentation of the occupational descriptions with and without the actual mentioning of role nouns. Essentially, the possibility of excluding the role nouns served as the experimental manipulation in the following Experiment 1B. Seventeen native French and 14 native English speaking participants who did not participate in the main experiment evaluated the constructed occupational descriptions on a 10-point rating scale as to how well they represented the role noun (1 = *the description is unlikely to represent the role noun* to and 10 = *very likely to represent the role noun*). Scales were reversed for half of the participants. There were four descriptions for each role noun, and the descriptions that received the highest evaluation and which corresponded in both languages were selected ($M = 8.51$, $SD = .84$).

The chosen experimental items were presented in two lists so that each occupational description and its corresponding role noun were combined with each block of pronouns. The experimental items were randomly mixed with 36 filler items to ensure that participants were reading different sentence structures during the experiment and to prevent them from associating gender as being part of the experimental manipulation. Similar to the experimental items, the filler items also consisted of three parts. There were three different clausal structures, 12 for each one (36 in total). The first structure illustrated a cause-effect situation beginning with *because* (i.e., *Because* reason...Effect), a time frame situation beginning with *before* (i.e., *Before* condition...Effect) or a conditional situation beginning with *if* (i.e., *If* condition.... Effect). The third part ended with a statement.

CAUSE-EFFECT FILLER:

- (1a) Etant donné que le bus s’est arrêté soudainement / Because the bus suddenly came to a stop,
- (1b) quelqu'un est tombé / someone fell over.
- (1c) C'était inattendu. / It was unexpected.

TIME-FRAME FILLER:

- (2a) Avant de partir en vacances, / Before leaving for our holiday,
- (2b) il faut arroser les fleurs. / we need to water the flowers.
- (2c) Autrement elles vont dépérir. / If not they will wilt away.

CONDITIONAL FILLER:

(3a) Si le sac est laissé sans surveillance, / If the bag is left unattended,

(3b) il va sûrement être volé. / it will surely be stolen.

(3c) Cette ville n'est pas très sûre. / This city is not very safe.

5.3.1.2.2 C-test

Four texts from Daller and Phelan's (2006) C-test were adapted to gauge participants' L2 English proficiency¹⁷. C-tests generally allow for an easy evaluation of a person's general language competency, that assesses L2 speakers' predictability for specific words and language constructions (Raatz & Klein-Braley, 1982). Participants were required to fill in a letter for each gap revealing their understanding for the obscured words within the text. The amount of correct restorations was then calculated and was considered as a measure of their L2 proficiency. The deletion procedure of the text was consistent with the method used in all C-tests, blocking out the second half of every other word except for the first and last sentences of each text.

5.3.1.2.3 Role noun check task

To ensure that participants actually understood the role noun cues, a list of all experimental role nouns in L2 English were given to the participants where they were instructed to provide a translation in L1 French.

5.3.1.3 Apparatus and procedure

Participants were individually tested in a quiet room. They were seated approximately 100 cm from a 1680 x 1050 pixel resolution monitor with stimuli presentation controlled by ExperimentBuilder software (SR Research Ltd.) One degree of visual angle subtended approximately 3.3 characters on average. The experiment was carried out using an Eyelink 1000 Desktop eye-tracking system (SR Research Ltd.) with a sampling rate of 500Hz. Viewing was monocular and only the dominant eye was tracked after each participant underwent a quick dominant eye test. Participants' heads were stabilized with a chin rest throughout the entirety of the eye-tracking portion of the experiment.

¹⁷ Although Sato et al. (2013) used an English C-test from Rahimi and Saadat (2005), Rahimi and Saadat's (2005) study only examined the construct validity of their C-test in terms of the competencies needed. In the present study, we adapted the C-test constructed by Daller and Phelan (2006) given that their study showed a significant validity of their tests in measuring L2 that was similar to that of the standardized TOEIC exam.

The experiment began with a nine-point calibration procedure for each participant. If the calibration procedure appeared to be inaccurate, this procedure was repeated until accurate calibration was obtained. Recalibrations were carried out between the items during the experimental session if the experimenter noticed tracking inaccuracies. A fixation point serving also as a calibration check appeared before each trial at the location where the stimuli were expected to appear. The experimenter triggered the presentation of each stimulus after the participant fixated on the fixation point and accepted proper calibration.

Experimental items were displayed in a way that the first part always appeared on one line at the top, the second part on the line just below, and the third part just below the second part. Participants were first presented with the first part, and were asked to press a button with their dominant forefinger to trigger the second part to appear. The third part was triggered similarly after the second part was read. Earlier presented sentences always remained on the screen while the later sentences appeared on the screen with it. One third of the passages were followed by a true / false comprehension question which participants responded with a *yes* or *no* button press. Participants were instructed to read naturally at normal reading speed and to answer the comprehension questions as quickly and accurately as possible.

A practice session involving four practice items were presented before the actual experimental session began to familiarize them with the procedures of the experiment. The experiment was divided into two sessions where participants read one list in French in the first session, and the other list in their L2 English in the second session. The administration of the second session was scheduled one week later as the task for the two sessions were similar. The language order of presentation was fixed given the paradigm involving occupational descriptions was new, and examining robust effects in the L1 French was necessary before comparing them with the weaker L2 English. The eye-tracking session in the second session was followed by the C-test and the role noun check task.

5.3.2 Results and discussion

5.3.2.1 Design and analysis

We report analyses based on the eye-tracking data only, and therefore participants' responses to the comprehension questions are omitted. Nonetheless, accuracy to the comprehension questions passed chance level (71% in L1 French and 70 % in L2 English).

The occupational description in the first part, the stereotypical role noun in the second part and the pronoun region in the third part were analyzed separately as primary regions of

interest. The pronoun region consisted of the pronoun itself, combined with the modal verb immediately following it (e.g., *she hoped*) as it was expected to be a probable spillover region and the pronoun itself was potentially too short to capture sufficient eye fixations. Short contiguous fixations less than 80 ms were merged together with larger fixations within one character and those remaining that were less than 80 ms were eliminated (5.3% in L1 French and 3.6% in L2 English).

To account for differences in region length, residual fixation times were calculated per participant (Trueswell, Tanenhaus, & Garnsey, 1994). This was done by first fitting a regression equation of the elicited fixation times against the region length (in number of letters) for each participant. Then, raw fixation times were subtracted from the predicted fixation times from the regression equation. Such a transformation of fixation times allows an examination of whether participants took relatively longer or shorter times to fixate on a specific region taking into account individual differences in reading speed. Negative fixation times indicate that the observed fixations times were faster than expected, whereas positive fixations times indicate that they were slower than expected.

For each critical region, we report four measurements. For what is referred to as early measures in the eye-tracking literature (e.g., Pickering, Frisson, McElree, & Traxler, 2004), we report first pass reading times and first fixation durations that are indicative of early processing during comprehension. These measures have been found as being sensitive to relatively early effects of language processing such as word frequency and general textual ambiguity (Pickering et al., 2004). First pass reading times refers to the sum of fixation durations in the region until it exits either to the left or to the right out of the region, and first fixation duration refers to the duration of the first fixation onto the target region. For measures that are known to reflect late processing (also known as late measures), we report regression path durations and total reading times. Regression path duration refers to the sum of fixation durations that occur on the target region until the region is exited to the right, including regressions to the left of the region whereas total reading time refers to the sum of all the fixations onto the region. While the discussions on which specific eye measurements represent what kind of processing are still ongoing, these reported measurements have been suggested by numerous psycholinguistic researchers (e.g., Carreiras & Clifton, 2004; Staub & Rayner, 2007) as useful indicators in psycholinguistic studies.

For the L1 data, the experiment was fully crossed with a 3 (gender stereotypicality of the role noun and description: female vs. male vs. neutral) X 2 (pronoun: she vs. he) within participant design. ANOVAs were conducted with both participants (F_1) and items (F_2) as

random factors and only those effects that were qualified by both by-participants and by-items analyses are reported. We report effects that were significant at alpha level of $p < .05$ and marginally significant at $p < .1$ (as in Duffy & Keir, 2004; Esaulova, Reali, & von Stockhausen, 2013; Irmen, 2007). Post-hoc contrasts were corrected with Bonferroni corrections.

Each occupational description region included different factors for the analyses of each eye-tracking measurement. At the description region, the analyses for first pass reading time and first fixation times included Stereotype as a within participant factor and Stereotype and Pronoun for the total reading time measurement. These effects were expected to result in non-significant outcomes given that there was no apparent reason to assume that reading a specific occupational description, in terms of stereotype, would result in differences at early levels of processing. The analyses were nonetheless conducted as a control measure to ensure that the materials were appropriate (e.g., not skewed for frequency).

The analyses for the role noun region for first pass reading time, first fixation times and regression path duration included Stereotype as a within participant factor, and Stereotype and Pronoun for the total reading time measurement. Finally the analyses for the pronoun region included both Stereotype and Pronoun as a within participant factor for all four measures. The respective by-item analyses included Pronoun as a within, and Stereotype as a between participant factor. Importantly, we did not expect effects associated to stereotype and pronoun resolution to surface at early measurements, but more so in the late measurements given that the activation and representation of gender requires elaborated and complex processes involving both stereotype and grammatical information

For the L2 English data, analyses took the same form as the L1 data, although only the experimental items that each participant reported to have understood were subject to analyses. Furthermore, participants' L2 proficiency was also taken into account for all regions of interest as a between participant factor for the F_1 analyses and a within participant factor for the F_2 analyses resulting in a 2 (Proficiency: high vs. low) X 3 (Stereotype: female vs. male vs. neutral) X 2 (Pronoun: she vs. he) mixed design. Proficiency groups were formed through a median split of the C-test scores. The high group consisted of 14 participants whose mean C-test score was 81%¹⁸ ($SD = 6.02$) and the low group consisted of 14 participants whose mean score was 63% ($SD = 9.88$). The two group differed significantly in their scores $F(1,$

¹⁸ The proficiency groupings for Experiment 2 consisted of groups with similar C-test score groupings. There were 13 participants in the high group whose C-test score were 80% and there were 16 participants in the low group whose C-test score were 69 %.

26) = 32.81, $p < .001$. Again, pronoun resolution effects were expected to surface on late measures.

5.3.2.2 L1 French data

5.3.2.2.1 Early measures

Occupational description region. No significant main or interaction effects were found in the first pass reading times or first fixation durations. Early measures are commonly predicted to show early stages of processing reflecting general syntactic difficulties or frequency effects (Pickering et al., 2004). As such, these results were expected, as there was no reason for the occupational descriptions to be read faster or slower at this point in reading. These findings support the view that the materials used in this study were reliable for experimental use.

Role noun region. As was the case for the occupational description region, no significant main or interaction effects were found in the first pass reading times or first fixation durations in the role noun region. Such null results were expected, further confirming that the role nouns were appropriate for experimental materials.

Pronoun region. While no effects emerged in the first fixation duration measurements, a significant Stereotype effect surfaced in the first pass reading times $F_1(2, 56) = 3.0, p < .1$; $F_2(2, 33) = 6.19, p < .05$. Further analyses suggested that fixations for the pronoun region were shorter when prior stereotypical information indicated neutral ($M = 10, p < .05$) than when it indicated male ($M = 107, p < .05$) stereotypicality. The female stereotype ($M_{\text{FPRT}} = 65$) lied in between the male and neutral stereotype information although much closer to the male stereotype. These effects suggested that lacking a specific gender implication facilitated early pronoun resolution process given that the readers did not have to assess the gender match between the pronoun gender and the previously implied gender. On the other hand, implying a probabilistic gender (i.e., stereotypically female or male) may impede pronoun resolution as readers would have had to disambiguate a general information mismatch.

5.3.2.2.2 Late measures

Occupational description region. No significant main or interaction effects were found in this region on total reading times. Participants did not revisit this region even to reassess the gender conveyed by this region after they continued on to reading the role noun region. These results demonstrate that role nouns suffice in providing gender information and a

reassessment of gender information conveyed by the occupational description may not be necessary.

Role noun region. Crucially for our study, a significant Stereotype effect emerged on total reading times in the role noun region $F_1(2, 58) = 15.34, p < .001$; $F_2(2, 33) = 3.82, p < .05$. Fixation times for the male stereotypical role nouns were the shortest ($M_{\text{Male}} = -183$) significantly differing from the female ($M_{\text{Female}} = -125, p < .05$) and neutral ($M_{\text{Neutral}} = -59, p < .01$) role nouns. The processing advantage for the male stereotype gender most likely resulted from the correspondence with the masculine grammatical marking of the role noun. It appears that the redundancy between stereotypical and grammatical information heightened the male gender in the representation facilitating information integration. In contrast, processing role nouns with the female stereotype was found to be more difficult to integrate given the clash between the stereotypical and grammatical information

Furthermore, total reading times revealed a main effect of Pronoun $F_1(1, 29) = 6.55, p < .05$; $F_1(1, 33) = 4.01, p \leq .05$. Participants faced greater difficulty processing the role noun when the following pronoun region specified a female ($M_{\text{TRT}} = -99$) than a male ($M_{\text{TRT}} = -146$) gender. These results showed that the masculine grammatical marking of the role noun had immediately determined a male gender which facilitated the processing of the gender congruent pronoun *il* [he] but impeded the representation of the gender incongruent *elle* [she]. As has been shown in past research, the male-specific interpretation of the masculine form was favored over the generic interpretation. No effects were found on regression path durations.

Pronoun region. An anticipated Pronoun effect was observed on total reading times with participants fixating longer on the pronoun region when the pronoun specified *elle* [she] ($M_{\text{TRT}} = -7$) than when it specified *il* [he] ($M_{\text{TRT}} = -101$). These effects suggested that the male-specific interpretation was facilitated over a possible generic interpretation of the role noun indicating that the masculine grammatical marking of the role noun had heightened the association to the male gender. Notably, this male-specific interpretation overruled any competing stereotypical information mentioned prior to the pronoun. Given that total reading times generally reflect information integration, this Pronoun effect appears to be reflecting (a) the assessment of the pronoun gender with prior stereotypical and grammatical information, and (b) an integration and updating (if needed) of gender information in the mental model.

5.3.2.3 L2 English data

5.3.2.3.1 Early measures

Occupational description region. There was an unexpected significant main effect of Stereotype reflected in first pass reading times $F_1(2, 50) = 10.85, p < .001$; $F_2(1, 33) = 5.42, p < .01$. While the male ($M_{\text{Male}} = 103, p < .001$) and female ($M_{\text{Female}} = 118, p < .001$) occupational descriptions did not differ significantly from each other, they were both fixated longer than the neutral occupational descriptions ($M_{\text{Neutral}} = -264$). Such Stereotype effects emerging on early measures were not anticipated, given that reading about an action associated to a certain gender should not in theory facilitate or hinder processing at such early stages of processing, and should be indicative of surface level effects. These results appear to be indicating the difficulty L2 English readers had with individual experimental items.

Role noun region. There were no significant main or interaction effects on either first pass reading times or first fixation durations at this region. These were anticipated results given that the processing effects of role nouns should not have emerged at such early stages of comprehension.

Pronoun region. Contrary to the results observed in the L1 French in which readers showed a facilitation effect for the neutral stereotype, no main or interaction effects emerged at this region on first pass reading times and first fixation durations. Gender stereotype information conveyed prior to the pronoun did not have an early effect in L2 pronoun resolution. In fact given the Stereotype effect emerging in the occupational description, it may be the case that readers did not properly understand them to influence pronoun resolution at early stages.

5.3.2.3.2 Late measures

Occupational description region. No main or interaction effects were found on total reading times for this region. Paralleling the results from the participants' L1 French, the results suggested that readers did not reassess the gender stereotypicality of the occupational descriptions even at later stages of processing after role nouns or pronouns were read. These may also suggest a lack of understanding of the presented occupational descriptions.

Role noun region. Although analyses for regression path durations did not indicate any significant effects, total reading times revealed a significant Proficiency effect, $F_1(1, 25) = 3.75; p < .1$; $F_2(1, 33) = 53.39, p < .001$. The role noun region was fixated longer by high than low proficient participants ($M_{\text{High}} = -84, M_{\text{Low}} = -237$). Intuitively, we would have

expected low proficient readers to take more time for basic decoding than high proficient readers. We interpret these results as indicating high proficient readers' commitment (i.e., invested time) to understanding the text. On the other hand, low proficient readers may have aimed to accomplish only general comprehension and may not have taken adequate time to properly construct their understanding regarding person references, possibly leading to poor comprehension.

Pronoun region. Interestingly, there was a significant Proficiency X Stereotype X Pronoun interaction observed in regression path durations $F_1(2, 48) = 3.08, p < .1$; $F_2(2, 33) = 2.83, p < .1$ (see Figures 5.1 and 5.2). Following Sato et al. (2013), this three-way interaction was initially anticipated where participants were expected to reveal processing tendencies conforming to gender biases that were characterized by the regularities of each language, with proficiency levels influencing the magnitude in which these characteristics would manifest. Specifically, high proficient L2 English readers were expected to show a stereotype-congruent effects, whereas low L2 English readers were expected to show a male-bias in processing as they would have relied on their dominant L1 French.

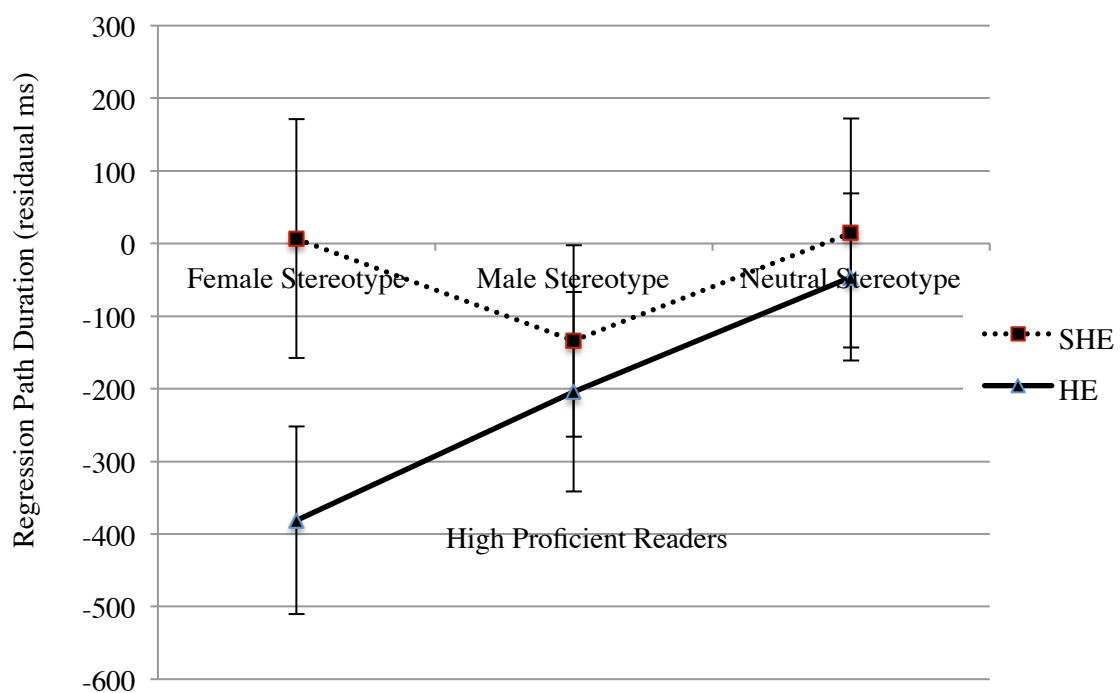


Figure 5.1 Effects of Pronoun at each level of Stereotype emerging in the pronoun region for high L2 readers. Error bars indicate standard errors.

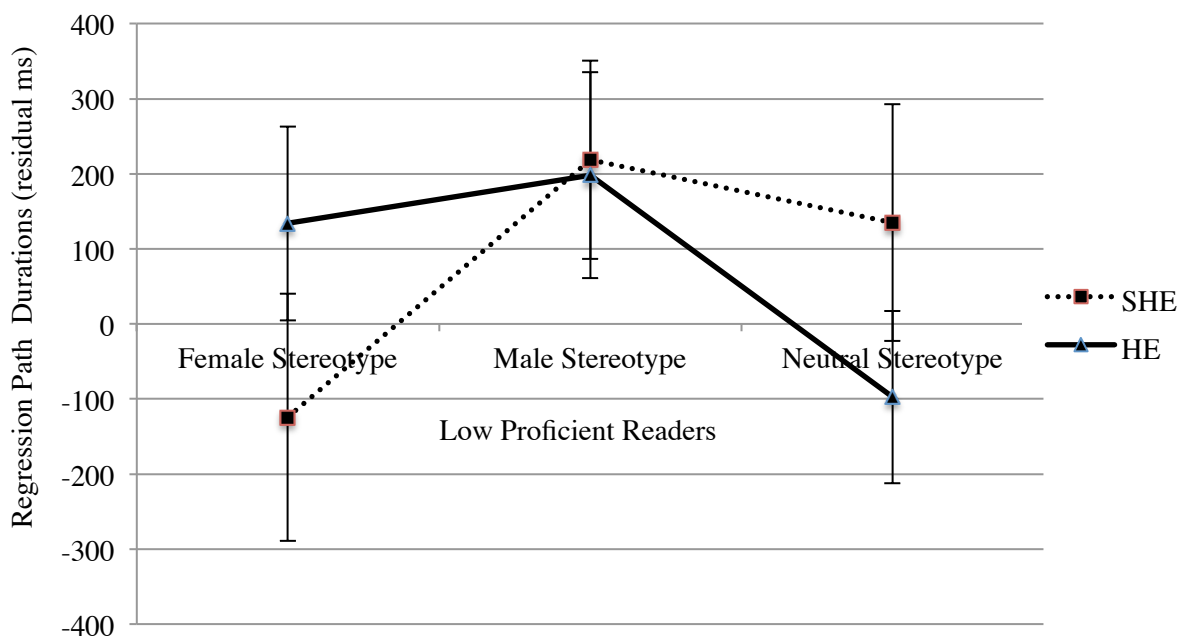


Figure 5.2 Effects of Pronoun at each level of Stereotype emerging in the pronoun region for low L2 readers. Error bars indicate standard errors.

Contrary to our predictions, post-hoc analyses for each proficiency group indicated that the effect of the interaction resulted from the differences between the pronoun *she* and *he* for the female stereotype among high proficient participants (Figure 5.1). Results revealed that high proficient participants fixations on the pronoun region were less when the specification of gender by the pronoun was a man ($M_{HE} = -381, p < .03$) than when it was a woman ($M_{SHE} = 8$) when prior information implied a female stereotype. Note that the female stereotype was expected to facilitate the comprehension of the pronoun *she* given a gender congruency-effect. The pronoun differences however did not emerge when prior information implied male or neutral stereotypical information. These results may suggest that the pronoun *he* was used as a default gender to apply to all person nouns as was the case in their L1 French. Alternatively, high proficient participants may have invested generally more time to represent *she* as can be seen for all stereotype conditions. Alternatively, the low proficient participants did not show any differences among the three stereotype conditions. Provided that significant Proficient effects emerging in regression path durations at the role noun region indicated that low proficient did not invest time to represent the described person reference, it is possible that low proficient participants did not pick up on the gender cues at all or were faced with a difficulty from the multiple gender cues. These results speak to the general

difficulty in resolving pronouns in the L2 as well as establishing a gender representation from various textual cues.

5.4 Experiment 1B

In Experiment 1A, the results suggested an activation of both stereotypical and grammatical gender information in L1, although ultimately grammatical gender information had a dominating effect in the representation process. In L2 English, although both high and low proficient readers struggled to represent gender information, high proficient readers invested more effort to represent the conveyed information.

To further examine the processes at hand, the activation of gender information was investigated in this experiment by examining whether gender information would still be activated in the absence of a gender categorical cue. Specifically, role noun cues that carry strong gender associations as part of their lexical representation were omitted, leaving only occupational descriptions (e.g., *shortening the trousers*) associated to a gender category to elicit gender inferences. Occupational descriptions were expected to suffice in specifying the representation without a need to further emphasize the gender by using a role noun. Nonetheless, the absence of the role noun was expected to moderately reduce the magnitude of its effects.

Moreover, unlike Experiment 1A which presented readers with an ambiguous masculine form in French – it could lead to interpretation difficulties if interpreted as being male-specific –, the present experiment did not pose any grammatical difficulties of the like.

5.4.1 Method

5.4.1.1 Participants

Twenty-nine L1 French-L2 English speakers (22 women, seven men) from the University of Fribourg, Switzerland participated in the study for course credits. None of them had participated in Experiment 1A.

5.4.1.2 Materials

5.4.1.2.1 Experimental stimuli

The experimental passages used in this experiment were the same as those used in Experiment 1A, with the exception that the second part including the role noun was eliminated. Thus participants were presented with the first part (French [a] and English [a])

including the description, followed by the third part (French [c], English [c]) that included the pronoun reference. The third part for each of the filler passages used in Experiment 1A were eliminated that resulted in a similar two-part structure to that of the experimental items, and employed as filler items.

5.4.1.2.2 C-test

The same C-tests used in Experiment 1A were employed.

5.4.1.2.3 Apparatus and procedure

The experiment was programmed and conducted as done in Experiment 1A. Given the absence of role nouns, the role noun check task as was conducted in Experiment 1A was omitted.

5.4.2 Results and discussion

5.4.2.1 Design and analysis

Apart from the analyses performed on the role noun, of course, the design and analyses procedures were identical to that of Experiment 1A. For the proficiency groups, the high group consisted of 13 participants whose mean C-test score was 80 ($SD = 3.67$) and the low group consisted of 16 participants whose mean score was 69 ($SD = 5.56$). The two groups differed significantly in their scores $F(1, 27) = 40.04, p < .001$. Participants responded to the comprehension questions above chance level (62% in L1 French and 82% in L2 English) and data filtering was conducted using the same procedure as Experiment 1A (3.6% in L1 French and 3.5% in L2 English).

5.4.2.2 L1 French data

5.4.2.2.1 Early measures

Occupational description region. As expected, neither first pass reading times nor first fixation durations presented reliable effects. Again, these non-significant effects were expected as they would have only indicated general processing difficulties.

Pronoun region. Crucial for our study, neither first fixation durations nor first pass reading times presented significant effects. These effects contrast those from Experiment 1A which revealed a significant Stereotype effect at the pronoun region (first pass reading times) indicating that a specific gender stereotype (i.e., female or male) hindered the initial stages of

pronoun resolution. Given the absence of a role noun prior to the pronoun in this experiment, these results suggest that a specific gender was not activated, or at least sufficiently, to influence the early processes when reading the pronoun.

5.4.2.2.2 Late measures

Occupational description region. Importantly, as indicated in Figure 5.3, the analyses revealed an anticipated Stereotype by Pronoun interaction emerging on total reading times $F_1(2, 52) = 3.07, p < .1$; $F_2(2, 33) = 2.68, p < .1$. These results mirrored those obtained by Real et al. (2015) who also found gender match-mismatch effects when presenting stereotypical descriptions followed by anaphoric pronouns in an eye-tracking study. Such results appeared to be mirroring current social dynamics in which a growing number of women are entering into male occupational roles (Diekmann & Eagly, 2000; Sczesny, Diekmann, & Twenge, 2007) and therefore may have facilitated the integration of female pronouns after female as well as male occupational descriptions. Nonetheless, on a descriptive level, the female occupational descriptions benefited the most when followed by the pronoun *elle* [she] reflecting a gender stereotype match.

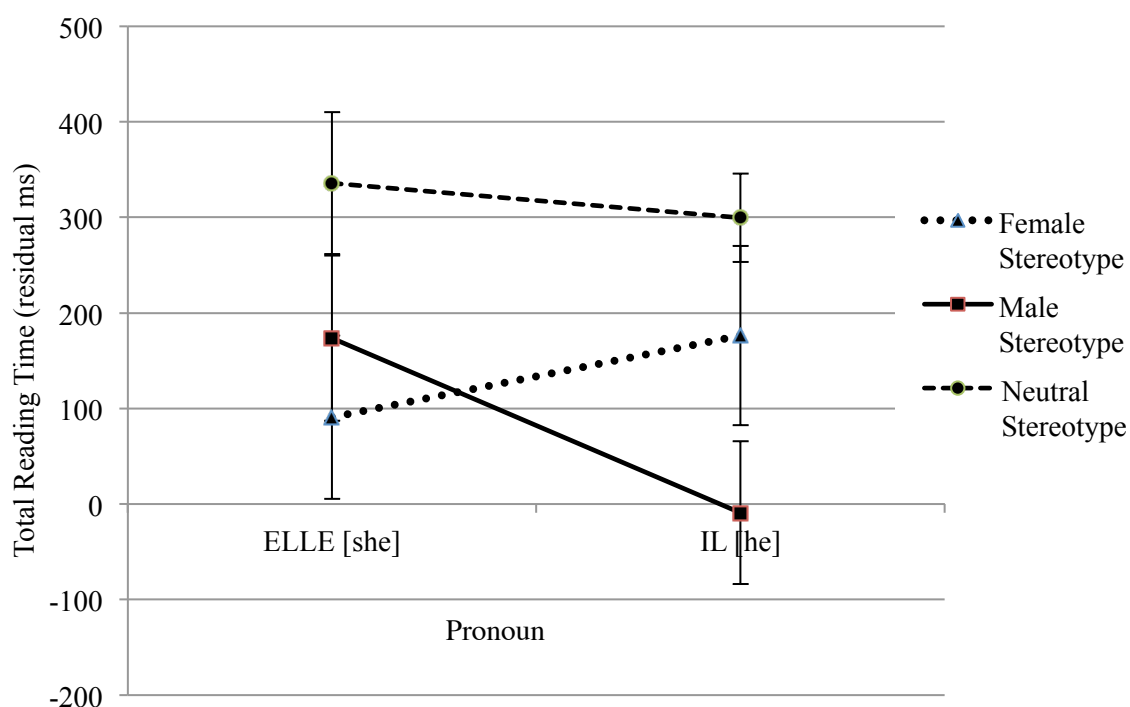


Figure 5.3 Effects of Stereotype at each level of Pronoun at the occupational description region in L1 French. Error bars indicate standard errors.

These results demonstrated that readers indeed associated occupational descriptions as being associated to a specific gender.

Pronoun Region. A significant Pronoun effect was observed on regression path durations $F_1(1, 26) = 14.64, p < .001$; $F_2(1, 33) = 22.35, p < .001$. When reading the pronoun region, the regions including *elle* [she] took longer time to process than when the pronoun region included *il* [he] ($M_{\text{She}} = 273, M_{\text{He}} = 110$). As regression path durations are considered as reflecting time that readers need to process information “to a sufficient degree that she is prepared to input new information” (Pickering et al., 2004), the Pronoun effect observed here can be differentiated from the Pronoun effect observed in Experiment 1A that surfaced in total reading times.

We believe that the Pronoun effect observed in Experiment 1A indicated that readers were evaluating the gender match between the male genders activated by the grammatical gender in the representation with that of the encountered pronoun gender. After this gender assessment, readers underwent an updating and integration process which was expected to be reflected at later stages of processing. However, provided that here, the similar Pronoun effect did not emerge in total reading times but rather surfaced in the regression path durations, it is likely that this Pronoun effect did not entail the similar processes associated to information updating and integration seen in Experiment 1A. In this experiment, encountering a pronoun reminded readers that a person reference needed to be represented in their mental model, which subsequently initiated a search for a probable person reference, and gender cue from prior text. Such an assumption would therefore suggest that occupational descriptions did not suffice in activating a specific gender or even a person reference. In this regard, only the presence of the pronoun itself played a role as a cue to initiate a gender search.

5.4.2.3 L2 English data

5.4.2.3.1 Early measures

Occupational description region. We observed a significant Stereotype effect emerging on first pass reading times $F_1(2, 54) = 16.55, p < .001$; $F_2(2, 33) = 4.28, p < .05$. Contrasts revealed that the effects paralleled the facilitation effect for neutral occupational descriptions on first pass reading times observed in Experiment 1A. The neutral occupational descriptions were fixated on significantly shorter ($M_{\text{Neutral}} = -188$) than female ($M_{\text{Female}} = -52, p < .001$) and male ($M_{\text{Male}} = 158, p < .001$) occupational descriptions. These effects reflect that our sample of L2 readers struggled to immediately comprehend the occupational descriptions, most likely resulting from their attention to access specific lexical items from their inventory and encoding processes. These surface level effects suggest that the L2 proficiency for the employed sample of participants may have been insufficient for the experimental task.

Pronoun region. In the analysis of first fixation durations, a Stereotype effect emerged $F_1(2, 52) = 3.97, p < .05$; $F_2(2, 33) = 5.04, p \leq .01$, revealing that the pronoun region was fixated longer when neutral occupational descriptions were presented prior to the pronoun, than when female occupational descriptions were presented ($M_{\text{Neutral}} = 22, M_{\text{Female}} = -4; p < .03$). The male occupational description ($M_{\text{Male}} = 5$) did not differ significantly from either stereotype. Early measures in the occupational description region indicated that L2 English readers took the least amount of time to process occupational descriptions with neutral stereotype, raising the possibility that they were not sufficiently processed at initial stages of processing. It thus makes sense that readers needed to spend the longest time on the pronoun region when neutral stereotype preceded this region, as they may not have been processed sufficiently.

5.4.2.3.2 Late measures

Occupational description region. Total reading times produced a highly significant Stereotype effect $F_1(2, 54) = 11.06, p < .001$; $F_1(2, 33) = 4.14, p < .05$. Further analyses revealed that male occupational descriptions were read significantly faster ($M_{\text{Male}} = -254$), than female ($M_{\text{Female}} = 16, p < .03$) and neutral occupational descriptions ($M_{\text{Neutral}} = 29, p < .03$) respectively. These results may be attributed to the commonly observed male-gender advantage, although Stereotype effects emerging in early measures have shown that the experimental stimuli may not have been appropriate for our sample of participants.

Pronoun region. No significant effects were observed for regression path durations or total reading times in this region. The lack of any results suggested that readers might not have attended to resolving the pronoun information as we would have intended.

5.5 Experimental Comparisons

In order to compare the magnitude between the occupational descriptions and role nouns as gender information cues, a combined analysis of both Experiment 1A and 1B was conducted on the occupational description and pronoun regions, with *Experiment* as an added between-participants factor. In the following analyses, we expected to observe a main effect of *Experiment* indicating longer fixations emerging in the pronoun region for Experiment 1B which consisted of reduced gender information cues than Experiment 1A. Such an effect would provide evidence to suggest the general difficulty in resolving the pronoun in the presence or absence of a strong gender-activating cue. Furthermore, an *Experiment* by *Stereotype* interaction emerging in either the occupational description region or pronoun

region in late measures were expected to reveal the different magnitudes of stereotype information that were conveyed from the different cues in each experiment. As the analyses were aimed at examining the difference between the two experiments, only the main effect of *Experiment* as well as interactions including this factor are reported.

5.5.1 L1 French data

5.5.1.1 Early measures

Occupational description region. No significant main or interaction effects with Experiments emerged on either first pass or first fixation durations. Early effects at this region were not expected to differ given that the conditions to read the occupational descriptions were the same in the two experiments.

Pronoun region. No main or interaction effects including Experiment emerged in the early measures at the Pronoun region. As such, both experiments showed similar patterns.

5.5.1.2 Late measures

Occupational description region. As expected, no main or interaction effects including Experiment emerged in the late measures at the occupational region.

Pronoun region. A highly anticipated main effect of Experiment emerged on regression path durations revealing that participants regressed substantially longer on the pronoun region in Experiment 1B ($M_{1B} = 191$) than in Experiment 1A ($M_{1A} = -95$) $F_1(1, 55) = 15.8, p < .001$; $F_2(1, 33) = 77.2, p < .001$. Overall, pronoun resolution was easier in Experiment 1A than in Experiment 1B given that the Experiment 1A provided role nouns that reconfirmed the specificity of the implied gender. In contrast, readers struggled to map the pronoun information given the weak gender activation by occupational descriptions in Experiment 1B.

There was also evidence for a significant Stereotype by Experiment interaction emerging in total reading times $F_1(2, 110) = 4.6, p \leq .01$; $F_2(2, 33) = 2.44, p < .1$ (see Figure 5.4). The interaction suggested that whereas the pronoun region in Experiment 1A did not differ in processing regardless of the stereotypicality denoted prior to the pronoun, the pronoun region in Experiment 1B was processed significantly faster when the aforementioned stereotype was male ($M_{\text{Male}} = -201, p < .001$) and female ($M_{\text{Female}} = -201, p < .001$) than when it was neutral ($M_{\text{Neutral}} = -85$).

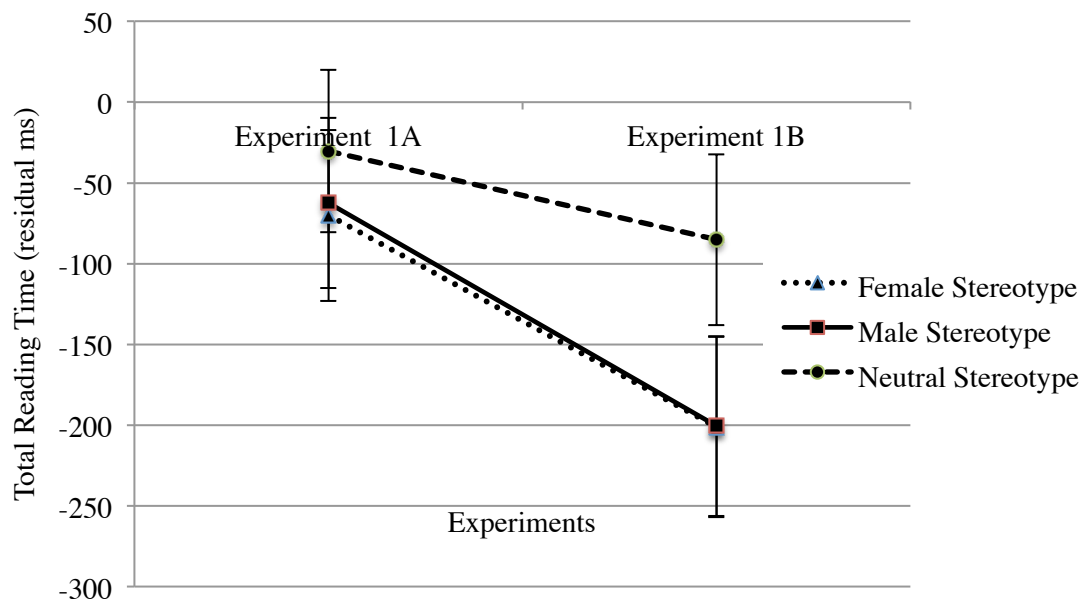


Figure 5.4 The effect of pronoun observed in L1 French according to each stereotype condition in both experiments. Error bars indicate standard errors.

These results suggested that presenting readers with role nouns in the masculine form (Experiment 1A) provided ample amount of gender information for readers to map the denoted person reference to the pronoun. In such a case, stereotype information did not affect the pronoun resolution process given that the grammatical information provided sufficient amount of information (essentially boosting the male gender) that could overrule stereotypical information. However in Experiment 1B, readers lacked the specific mentioning of a person reference until they encountered a pronoun. When the search for cues was initiated, mapping the pronoun to references that were associated to a specific gender (i.e., female and male) was more accelerated than those without a specific gender association (i.e., neutral stereotype).

5.5.2 L2 English data

5.5.2.1 Early measures

Occupational description region. A Stereotype by Experiment interaction emerged on first fixation durations $F_1(2, 104) = 3.78, p \leq .05$; $F_2(2, 33) = 2.68, p < .1$ (see Figure 5.5). While processing speed for male ($M_{1A} = 14, M_{1B} = 4$; *ns.*) and neutral ($M_{1A} = 1, M_{1B} = 11$; *ns.*) stereotype information did not differ between the experiments, the female stereotype was processed significantly faster in Experiment 1B than 1A ($M_{1A} = 18, M_{1B} = -5$; $p < .03$). Interactions consisting of early Stereotype effects are most likely associated to specific experimental items. Given that there were longer fixations for Experiment 1A in the female

stereotype condition this may suggest that the correspondence of gender information for the female stereotype may not have been strong as the other two stereotype information.

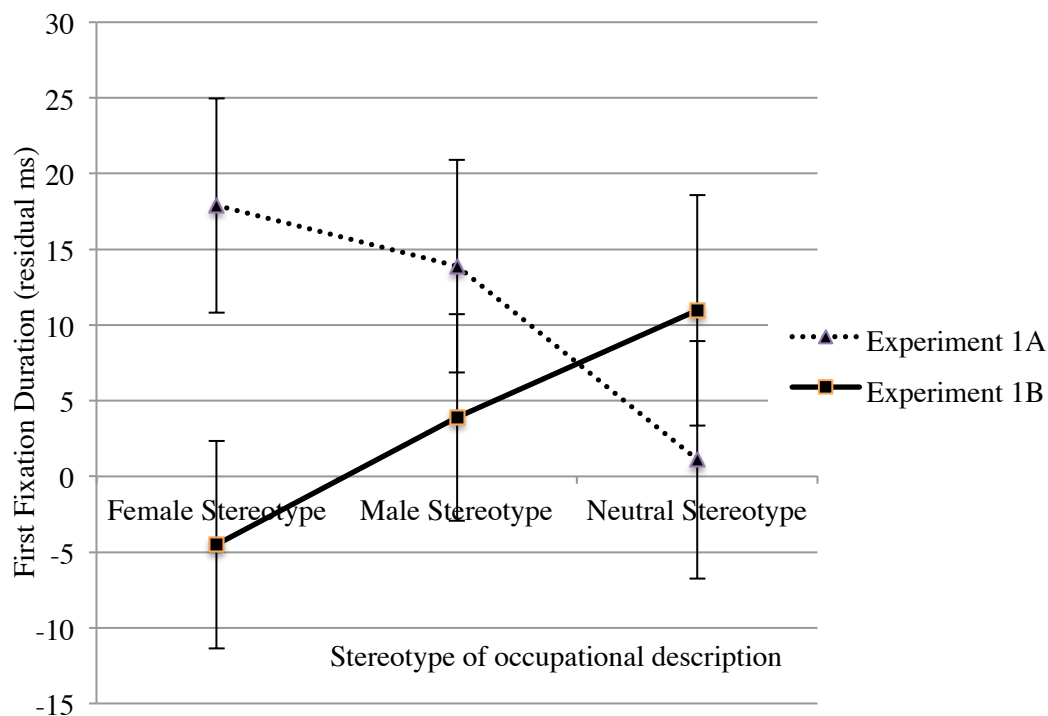


Figure 5.5 Stereotype effects for each L2 Experiment reflected in first fixation durations. Error bars indicate standard errors.

Pronoun region. A reliable Experiment effect was observed on both first pass reading times $F_1(1, 50) = 11.68, p \leq .001$; $F_2(1, 33) = 143.22, p < .001$ and first fixation durations $F_1(1, 50) = 3.54, p < .01$; $F_2(1, 33) = 29.46, p < .001$. These early pronoun effects suggested that participants fixated significantly longer on the pronoun region in Experiment 1B ($M_{\text{FPRT}} = 343, M_{\text{FFD}} = 7$) than Experiment 1A ($M_{\text{FPRT}} = 118, M_{\text{FFD}} = -11$). These results suggested a general ease in disambiguating gender information at very early stages of pronoun processing for L2 readers in Experiment 1A, provided that the role nouns were able to hint at the implied gender stereotype.

5.5.2.2 Late measures

Occupational description region. No main or interaction effects including Experiment emerged in the occupational description region as expected.

Pronoun region. Both regression path durations $F_1(1, 50) = 16.55, p < .001$; $F_2(1, 33) = 116.76, p < .001$ and total reading times $F_1(1, 52) = 5.44, p < .05$; $F_2(1, 33) = 153.26, p < .001$, revealed significant Experiment effects. Consistent with the early measures, the

directionality for regression path durations revealed longer regressions for Experiment 1B into the pronoun region ($M_{1B} = -23$) than 1A ($M_{1A} = 299$). Given that the gender information conveyed by occupational descriptions may not have been integrated by our L2 participant sample, although role nouns were understood, it makes sense that pronoun resolution took longer in Experiment 1B than 1A, where gender inferences could have been generated based on the role nouns. Interestingly, given that the ultimate information integration process was nonetheless complex with readers undergoing a gender-match process, total reading times were found to be longer in Experiment 1A than 1B ($M_{1A} = 12$; $M_{1B} = -170$).

5.6 General discussion

Research has provided ample evidence to suggest that inferences about a person's gender are generated when reading stereotypically biased role nouns that may be influenced by conceptual and/or grammatical constraints (Braun, Sczesny, & Stahlberg, 2005; Carreiras et al., 1996; Gygax et al., 2008). In the two experiments presented in this study, we investigated the magnitude in which role nouns and occupational descriptions would impact the representation of gender information during reading.

In Experiment 1A, we examined the activation of gender information by presenting participants with occupational descriptions associated to a specific gender followed by a gender emphasizing role noun in readers' L1 French and L2 English. While role nouns were both conceptually associated to either gender and marked with a masculine grammatical marking in French, they were only conceptually associated to either gender in English. In the following Experiment 1B, the gender-biased role nouns were eliminated to remove the strong gender emphasis from both languages. Readers were thus faced with more subtle gender cues which consisted of only partial gender features for representational purposes than the strong role noun cues in Experiment 1A.

5.6.1 L1 French results

The results of Experiment 1A and 1B lend support to the hypothesis that role nouns and occupational descriptions provide gender cues that instigate varying magnitudes of gender information. While in the former Experiment 1A, there was evidence to suggest a strong activation and an intricate influence of both stereotype and grammatical information during pronoun resolution, Experiment 1B suggested that readers instigated a search for

gender information only when readers came across a pronoun and became aware that a possible antecedent needed to be represented in their mental models.

A crucial late Pronoun effect emerging at the pronoun region (reflected in total reading times) was found when readers read both role nouns and occupational descriptions (Experiment 1A). The effect showed that pronoun resolution was delayed for the female pronoun *elle* [she] compared to the male pronoun *il* [he], most likely resulting from the fact that the male gender had already been activated in the readers' mental models. If the gender conveyed by the pronoun was at odds (i.e., *elle*) with the male gender previously represented in the mental model, it essentially resulted in longer fixation times to compensate and update the mismatching information. The most obvious justification for the activated male gender was accredited to the presence of the masculine grammatical marking on the role noun, which heightened the association to the male gender. As was the case with prior studies investigating the interpretation of the masculine grammatical form (e.g., Braun et al., 2005; Gygax et al., 2008), readers took the masculine form as referring specifically to the male gender despite its possible generic denotation, which overrode any competing stereotypical information conveyed by the role noun. These results pointed to the prominence of grammatical influences that persisted irrespective of competing stereotype information, further promoting male-dominant representations.

However, notwithstanding the ultimate impact of grammatical information, stereotypical information was still found to influence pronoun resolution. The pronoun region revealed an immediate activation of Stereotype (reflected in first pass reading times), which crucially affected gender representation at early stages of pronoun processing. The activation of specific stereotypical genders (i.e., male or female) was found to delay pronoun resolution, whereas the absence of gender information (i.e., neutral stereotype) accelerated the process. Assuming that early effects reflected in first pass reading times are commonly attributed to disambiguation of general information mismatches (Pickering et al., 2004), it was reasonable to find neutral stereotype as having the greatest processing advantage as readers did not have to map the gender conveyed by the pronoun with the aforementioned stereotype information. These effects are consistent with the findings offered by Irmen (2007) who also regarded stereotype information as impacting gender representation at early stages of processing although finding an ultimate override of grammatical information.

Moreover, stereotype influences also surfaced in late measures (total reading times) at the role noun region. If role nouns were supplemented by a gender correspondence between the grammatical marking and its stereotype (i.e., masculine form and male stereotype),

information integration was accelerated substantially. Alternatively, when the masculine grammatical gender was complemented by female stereotype information, the violation of gender between the two sources of information rendered gender integration to be most difficult. Together these results lend support to the findings reported in production studies by Vigliocco and Franck (2001) suggesting that when both conceptual and grammatical information are congruent, they may strengthen syntactic information processing whereas contradicting information may have a detrimental impact. We therefore complement Irmen's (2007) conclusions by proposing that while affecting initial stages of gender representation, stereotype influence may in fact carry on into later stages of processing, essentially functioning together with grammatical information.

Notably, although a pronoun effect also emerged in late measures when reading occupational descriptions (Experiment 1B), in the absence of strong role noun cues, these effects were found to be less substantial than the effects found in Experiment 1A. Such an account was supported by the analyses comparing the two experiments as well as in the pronoun effect emerging in Experiment 1B importantly found only on regression path durations but not on total reading times. Because regression path durations merely indicate processing representative of a “sufficient degree” of comprehension, it is possible that the readers did not actively commit to resolving the pronoun immediately especially taking gender into consideration, as was seen in Experiment 1A. In fact, contrary to our hypothesis, it is possible that readers had not even activated a specific gender upon reading the pronoun, but rather encountering the gender specific pronoun region functioned as a reminder to reexamine prior text information for any gender cues that could assist in the representation of person references. The absence of such an automatic gender activation would explain why no early stereotype effects were found in the pronoun region of Experiment 1B, and more importantly, by the crucial stereotype by pronoun interaction emerging at the occupational description region in total reading times. Lacking these early stereotype effects at the pronoun region in Experiment 1B suggested that the presentation of occupational descriptions alone did not suffice to activate a specific gender stereotype, although the interaction indicated that when readers actively searched for gender cues, the occupational descriptions were nonetheless associated to a specific gender. A gender congruency between the stereotype of occupational descriptions and pronoun gender facilitated the representation process than when there was a violation between the two. Such results concurred those found by Reali et al. (2015) who also found the impact of stereotype effects triggered by gender typical descriptions in their German study.

Our findings clearly point out that the activation of gender information was less substantial when reading only occupational descriptions in the L1 French than when these occupational descriptions were presented in conjunction with a gender emphasizing role noun. Analyses comparing the two experiments showed that readers required more time to process the pronoun for Experiment 1B than 1A given that readers did not have strong enough gender information activated prior to reading the pronoun. Particularly for Experiment 1B, lacking gender implications instigated by gender categorical labels was found to complicate pronoun resolution even more.

These results are in line with the initial assumptions that role noun cues consist of whole features of person information accompanied by formal grammatical information that may also contribute in representation processes. These conditions provide for stronger information cues that instigate strong activation of both stereotypical and grammatical gender information. Particularly, stereotype information was found to come into effect at very early stages of processing and persisted into later stages of processing.

While we did not observe any early influences of grammatical information in Experiment 1A (i.e., early Pronoun effect at the Pronoun region), we propose that we cannot completely isolate the effects of grammatical gender at early stages of processing. One reason for this can be attributed to our experimental paradigm employing a pronoun resolution task. Even if readers had activated grammatical gender information immediately when reading the pronoun, the relevant information (e.g., generic interpretation, male-specific interpretation) contained in grammatical gender that was necessary to resolve the pronoun would have had to be assessed and chosen. It is likely that the assessment of such information would only have manifested at later stages of pronoun processing. Further research may consider isolating the effects to confirm the time course in which grammatical gender information is resolved and activated.

A final issue that should also be mentioned here is the advantage for the pronoun *il* [he] found in both experiments. While it is claimed that the grammatical marking on the role noun in Experiment 1A brought about the specification of the male gender facilitating the male pronoun, the general advantage seen in the early stages of processing (observed in the experimental comparisons) could be attributed to the fact that *il* [he] is used considerably more frequently than *elle* (New et al., 2004). The difference in the use of the two pronoun forms arises from the fact the pronoun *elle* can only be used exclusively to refer to women and feminine nouns (i.e., *it* in English), although *il* is commonly used to refer to men, masculine nouns and as a generic sense. The multiple functions that the male pronoun entails

most likely triggered an immediate and early processing advantage for *il* [he]. Another possibility for the *il* [he] advantage can be associated to the common difficulty for people to represent women given that social norms presuppose a *people = male* view (Hamilton, 1991; Silveira, 1980).

5.6.2 L2 English results

With regard to our L2 English results, we initially expected L2 readers to conform to gender stereotype-consistent tendencies given that stereotypical knowledge about the world dictates gender representations in L1 English. Importantly, the impact of these effects were expected to be affected by readers' L2 proficiency provided that past research has reported that the magnitude of these shifts were dependent on readers' L2 proficiency (Sato et al., 2013). As competencies in L2 advanced, representation and comprehension mechanisms were expected to shift to more native-like tendencies of the target language given the greater influence of the language. On the other hand, low proficiency was expected to result in a greater reliance of the L1 French, consequently leading to representations that were male biased. These hypotheses were met with mixed results, essentially revealing a general difficulty in L2 pronoun resolution as well as issues with our participant sample.

As revealed by the analyses for experimental comparisons, participants struggled to integrate the implied gender information in Experiment 1B than in 1A with fixations (apart from the total reading time) in the pronoun regions being longer in 1B than in 1A. Furthermore, an absence of the Pronoun or a Pronoun by Stereotype interaction as seen in Experiment 1B in the occupational description region or pronoun region suggested that a probabilistic gender was not activated during reading. Although we would have assumed these results as showing the reduced magnitude in which occupational descriptions activated gender information, they touched on several considerable issues.

To begin with, the difference of the impact between the role nouns and occupational descriptions was difficult to determine as the results alluded to the possibility that L2 readers had not sufficiently comprehended the presented occupational descriptions. Stereotype effects emerging in first pass reading times were found at the occupational description region in both experiments demonstrating that certain stereotype items posed different levels of comprehension difficulty at early stages of processing as opposed to role nouns that did not exhibit such early Stereotype effects. Assuming that early effects reflect difficulty arising from general textual ambiguity, these effects were highly unanticipated. Indeed, although we were attentive at verifying whether the participants had wholly understood the presented role

noun through an off-line verification task, occupational descriptions were not tested for in the same manner.

Nonetheless we argue that even if participants had properly understood the presented occupational descriptions, they would not have provided for strong gender cues that would have instigated an immediate activation of gender. Such a view is supported by the fact that in the L2 English, even role nouns that contain whole categorical features did not provide for a strong gender cue that impacted participants' representations. Contrary to the L1 French where the impact of the role noun stereotype immediately activated gender information affecting pronoun resolution at early stages of processing, no similar effects emerged in the pronoun region to suggest that a specific gender had already been activated. Crucially, a Proficiency by Stereotype by Pronoun interaction emerged in the pronoun region (reflected in regression path durations) suggesting that even high proficient participants relied on their L1 French processing tendencies while reading in their L2 English, that resulted in a male-biased representation. No stereotype advantage was seen when processing the pronoun *he*, therefore confirming the general facilitation for the male pronoun. Interestingly, the pronoun *she* took longer time to process than *he* when prior information conveyed female stereotype information which we interpreted as indicating high proficient readers' efforts to represent the conveyed gender. The fact that the same high proficient readers took more time to process the role noun containing the most gender-relevant information also contests to their elevated motivation to properly represent textual information.

We believe that the lack of these gender effects as well as L2 proficiency effects stemmed from our participant sample and that the selected participants had not developed a sufficient level of L2 proficiency to engage in proper representation process required in the experimental task. We stress here that the task required intricate anaphoric resolution processes that required an encounter of the pronoun in order to initiate the antecedent search even for a native speaker. Specifically, our L2 readers needed to search for an antecedent in the absence of explicit person references such as role nouns. We also argue that gender representation processes consist of generating elaborative inferences that are normally made to enrich textual information, and is not a requisite to maintain text coherence. In this regard, our sample of participants may not have had the scope to generate such high level inferences. Although we cannot make a direct comparison of participant samples to that of Sato et al., (2013) due to the different C-tests used, our participants in the high group (Experiment 1A: 80%; Experiment 1B: 81%) scored roughly the same as their high group (82%) in a C-test with more or less the same text difficulty. The fact that our participants did not reach a level

where representational changes could be observed further confirms the relative intricacy of the processes involved in extracting gender information from different textual cues. Additionally for L2 readers, inferring a person reference in the absence of an explicit person reference was an intricate reading task that imposed general comprehension difficulties.

5.7 Conclusion

We have provided evidence to suggest that the magnitudes in which gender information is conveyed differs substantially according to whether (i) readers are presented with occupational descriptions that are associated to a specific gender or whether (ii) the occupational descriptions are also supplemented with a gender emphasizing role noun that carries whole categorical information in the lexical representation. Our results suggested that in both L1 French and L2 English, role nouns evoke elaborative inferences although the encounter of the occupational description may impose different levels of difficulty in each L1 and L2. In the L1 French, occupational descriptions did not instigate immediate effects but ultimately provided a source of information that enriched the readers' mental model. In the L2 English, the representation of gender was found to be difficult especially in the absence of whole categorical cues that may require a certain level of L2 proficiency before the representation process is initiated.

Chapter 6 : Study IV

Altering male-dominant representations: A study on nominalized adjectives and participles in first and second language German

The study presented in this chapter has been submitted for publication as:

Sato, S., Gabriel, U., & Gygax, P. M. (2015). Altering male-dominant representations: A study on nominalized adjectives and participles in first and second language German.

6.1 Abstract

Studies I – IV demonstrated that role nouns provide for cues that initiate strong gender activations. In fact, even when role noun cues, including whole features of categorical information, were omitted (Study IV), a specific gender was found to be represented with moderate degrees. Following these studies, the present study was aimed at examining forms that may (depending on the recency of their acquisition) attenuate these gender biases.

Specifically, the present study examined whether the recently introduced gender-neutral forms of nominalized adjectives and participles in German provide references that induce more balanced representations. We used cross-linguistic differences as a means to illustrate the flexibility of the gender representation system and investigated both native and non-native (French-German bilinguals) speakers of German. Although a masculine bias persisted when participants read role nouns in the masculine plural form, the study suggests that, for native speakers, the usage of nominalized forms does attenuate this male bias. However, results from our non-native participants suggest that although nominalized forms should not induce any biases, there is still an inherent male bias in their native language (French). The results of the study provide further support for the use of gender-neutral language.

6.2 Introduction

Research in social psychology and psycholinguistics has revealed striking differences in language-based gender biases that act as a vehicle for communicating and perpetuating

gender disparity (Prewitt-Freilino, Caswell, & Laakso, 2012; Stahlberg, Braun, Irmen, & Sczesny, 2007; Wasserman & Weseley, 2009). Especially for grammatical gender languages like German, French and Spanish, gender information embedded in the grammatical system has been found to activate a specific gender during language comprehension processes, even in the case of word forms that are not intended to refer to any gender (Stahlberg et al., 2007). In particular, a substantial amount of research has shown that individuals tend to associate masculine forms of words with a male-specific gender, even when the word is intended to refer to both genders (Gygax, Gabriel, Sarasin, Oakhill, & Garnham, 2008; Irmen, 2007; Irmen & Kurovskaja, 2010; Stahlberg, Sczesny, & Braun, 2001). Furthermore, this general facilitation to activate the male-specific gender has been found to override stereotypical gender information (e.g., Gygax, Gabriel, Lévy, Pool, Grivel, & Pedrazzini, 2012; Gygax et al., 2008). With an increasing societal demand for an equal representation of the sexes, changes in language use (i.e., gender-neutral language) have been encouraged in an attempt to make the two sexes equally visible. Paradoxically however, gender-fair language has also been criticized as distorting language, with research suggesting that its production requires more effort and language competence (e.g., Koeser, Kuhn, & Sczesny, 2014; Kuhn & Gabriel, 2013) than the more common generic masculine forms.

In the present paper, we aimed to demonstrate that using gender-neutral language might possibly be very efficient in alleviating male biases in German despite its strength, as shown in past research. We explored how the use of particular grammatical forms can directly tap and eventually influence our mental representations by comparing grammatically masculine and nominalized forms in German. Whereas the former masculine form (e.g., *Die Käufer*_{Masculine} [the buyers]) has been found to trigger male-biased representations despite its possible generic sense, the latter nominalized form refers to nouns that are derived from adjectives and participles (e.g., *Die Konsumierenden* [those that consume]), and are thus grammatically neutral and gender-unbiased. Although there are similar on-going debates on the use of gender-neutral forms for most grammatical gender languages, nominalized forms are particularly interesting in German (in contrast with other languages, as discussed later), as they are truly neutral (i.e., completely dissociated from semantic and grammatical gender). Considering these cross-linguistic differences as well as the limited usage of gender-neutral forms, we assessed whether these forms could be representationally effective in neutralizing activated gender biases.

In particular, a bilingual population was employed in our study to observe whether language does indeed influence the representation process. If we observe a difference between

the representations according to the bilinguals' linguistic domain, it should provide us with convincing evidence as to the extent the language being used for comprehension dominates representation tendencies. Specifically, bilinguals who have acquired two grammatical gender language systems that differ in gender-neutral forms (i.e., French and German) were employed for the study. The idea that specific languages influence mental representations is linked to the thinking-for-speaking hypothesis (Slobin, 2003). This hypothesis suggests that engaging in language processing directs readers' attention to certain attributes that are grammaticized within a language (Slobin, 2003). In terms of gender representations, if bilinguals switch to a language that specifically integrates certain gender information in its grammar, they would need to attend to this information for the purpose of basic comprehension. We argue that as nominalized forms in German have no grammatical association to either gender, they should not activate any gender during reading. Similarly, non-native German speakers (i.e., native French speakers) should be able to effectively alter their representations when reading nominalization forms in German even if a matching form does not exist in their first language (L1). In the following, we review research associated with gender representation during language comprehension and broaden our case for the usage of nominalized forms in our study.

6.2.1 Grammatical gender languages and gender representation

In grammatical gender languages such as German, French or Spanish, gender is a fundamental grammatical element that exerts an influence at different levels of language processing. In these languages, most regular nouns are attributed to an arbitrary gender class (e.g., masculine, feminine, and sometimes neutral) that is not directly associated with its meaning (e.g., German: *eine Tür*_{Feminine} [a door]; *ein Schlüssel*_{Masculine} [a key]). In contrast, most person nouns have a gender attribution directly linked to the biological sex of the person taking the position (e.g., German: *ein Schüler* [a male student]; *eine Schülerin* [a female student]). While this latter rule appears rather well-defined, the markedness of the two gender forms differs for female and male references, often causing asymmetry when interpreting them. The feminine form refers solely to a woman or a group of women, showing a direct relation between its surface form and its meaning. In comparison, the masculine form can refer to a man, or a group of men, but also to a group composed of both men and women (i.e., generic), or is applied in circumstances when the sex can be overlooked. The generic interpretation of the masculine form is thus inconsistent with the association of its surface form. The two possible interpretations of the masculine form posit a certain level of

ambiguity (e.g., Gabriel & Gygax, 2008; Gygax et al., 2008; Irmen, 2007; Lévy, Gygax, & Gabriel, 2014), which has been illustrated in existing research in German (e.g., Braun, Gottburgsen, Sczesny, & Stahlberg, 1998; Braun, Sczesny, & Stahlberg, 2005; Irmen, 2007), French (e.g., Brauer & Landry, 2008; Gygax et al., 2008; Lévy et al., 2014), Norwegian (Gabriel & Gygax, 2008) and Spanish (Nissen, 2013).

As suggested by the mental models theory of language comprehension (Garnham & Oakhill, 1996, Johnson-Laird, 1983), readers will integrate explicitly-provided morphological information (i.e., grammatical gender information) and their individual world knowledge about gender biases during gender representation. Inferences about the probable gender will be generated irrespective of whether it is essential for comprehension or not. Research shows that generic inferences when reading the masculine form (e.g., *Geburtshelfer* [birth attendants]) are particularly difficult to activate, and is commonly interpreted as referring more to men than to women (e.g., Braun et al., 2005; Gygax, et al., 2008; Irmen & Knoll, 1999; Stahlberg et al., 2001). Association to the male gender is heightened through its morphological form, and its generic sense becomes more difficult to activate. Note that studies on anaphor resolution have reported the effects of stereotype when investigating the interaction between stereotype and grammatical gender information in grammatical gender languages, yet only when looking at particular processing time windows (e.g., Esaulova, Reali, & von Stockhausen, 2013; Irmen, Holt, & Weisbrod, 2010; Irmen & Schumann, 2011; Reali, Esaulova, & Von Stockhausen, 2015). This strongly suggests that although most studies have reported strong male associations from specific morphological associations, stereotype information is still activated, but is most likely overridden.

In sum, these studies demonstrate that grammatical gender impacts the interpretation of human referents by increasing the saliency of a specific gender. These representation mechanisms enhanced by grammatical cues provide a context for investigating the impact of gender-neutral language, which is often encouraged by formal institutions (e.g., American Psychological Association, 2001; Duden, 2009). Presenting gender-neutral forms that are morphologically unrelated to either gender should serve to attenuate the gender-emphasizing effect, leading to a more neutralized representation. If the process of understanding language guides readers' attention to focus on specific information such as gender, then we should also see an impact of nominalization forms on gender representations for non-native speakers in their second language (L2), even if their representation preferences are well characterized by their L1.

6.2.2 Nominalized forms as gender-neutral language in German

The implementation of gender-fair language aims to establish a symmetry between the sexes by making women more linguistically and grammatically visible through means such as feminization (e.g., pair forms: *Politikerinnen*_{Feminine} und *Politiker*_{Masculine} [female politicians and male politicians]; splitting forms: *Politiker/innen*; adapting capital I to emphasize the feminine in German - *PolitikerInnen*) or by means of neutralizing the expression of both sexes (e.g., neuter nouns: *Individuum* [individual]; nominalized forms : *Alten* [elderly]) (Duden, 2009). Psycholinguistic literature on the influences of gender-fair language use has focused mostly on the effects of feminization (e.g., Bem & Bem, 1973; Chatard, Guimont, & Martinot, 2005; Stout & Dasgupta, 2011; Vervecken, Hannover, & Wolter, 2013), yet research addressing the impact of gender-neutral language usage has not yet clearly shown, or defined, its individual effects on mental representations of gender. The present study aims to fill this gap by examining a specific case of gender-fair language – neutralized forms in German through nominalized forms – and its impact upon mental representations of gender.

In German, nominalized nouns are commonly derived from adjectives and participles, and although they maintain grammatical gender in their singular form, many plural forms are neutralized and lose grammatical gender completely (e.g., *die Alten* [the elderly]-*alt* [old_{adjective}], *die Studierenden* [the students] – *studieren* [to study_{verb}]). As these forms do not originate from nouns that carry grammatical gender but from gender-unmarked adjectives and participles, they are considered entirely unassociated with any gender and should not in theory activate any gender connotations. While nominalization forms offer an attractive possibility to decrease male biases, studies on neutral forms nonetheless remain sparse. For example, Braun et al. (1998), in a series of experiments, presented participants with a text about a fictitious meeting of a scientific association written with either masculine generic forms (e.g., *die Geophysiker* [the geophysicists]), pair forms (e.g., *Geophysikerinnen und Geophysiker* [female and male geophysicists]) or neutral forms (composed of nominalized forms, e.g., *die wissenschaftlich Tätigen* [the people active in science], as well as other neutralizing forms such as *die Geophysik* [the field of geophysics]). The authors found that the usage of pair forms increased readers' estimation of female attendees to the meeting, yet there was no clear evidence of neutral forms having any impact. Similar findings were reported by Stahlberg et al. (2001) who compared the same three categories and asked participants to name their favorite heroes, musicians, athletes and singers.

More relevant for our study, Irmen and Roßberg (2004, Exp. 2) investigated the relationship between grammatically neutral forms and gender stereotypical information. They examined reading times of sentences with specific gender continuations (e.g., female continuation: *When going out they prefer for example a dress.*) that were preceded by a sentence including either pair forms (e.g., *Male and female soldiers only wear uniforms when on duty.*) or gender unmarked nominalized forms in German. Nominalized present participles with neutral (e.g., *Lehrende* [those who teach]), female (e.g., *Alleinerziehende* [those who raise a child alone]) or masculine stereotype (e.g., *Vorstandsvorsitzende* [those who chair the steering committee]) were presented. Reading times for nominalized forms indicated that a match between male stereotypical role nouns and male continuations showed accelerated reading times, while a match between female stereotypical role nouns and female continuations did not. Interestingly, when role nouns did not have a gender stereotype, participants not only showed accelerated reading times for continuations that were neutral (i.e., did not specify gender), but also for male continuations, at least when compared to female continuations. The authors interpreted these results in terms of neutral grammatical forms that also generated male-biased representations, based on the idea that *people = male* (Hamilton, 1991; Silveira, 1980). While this is a plausible explanation, their study was published in 2004 when nominalized forms had just started to emerge in German. The lack of exposure to these forms could have been the reason for why there was no substantial effect in mitigating the masculine bias.

6.2.3 Cross-linguistic issues on nominalized forms and gender representation

In the present study, we explored the impact that language has on the construction of gender representation by examining the neutralizing effects of German nominalizations among native and non-native speakers of German (i.e., French-German bilinguals). French-German bilinguals provide an interesting contrast to those of native German speakers given that nominalization forms in their L1 French differ from that of German forms. In contrast to German, nominalization forms in French carry a specific grammatical gender (e.g., *étudiante*_{feminine} vs. *étudiant*_{masculine} [a female vs. a male person who studies]). While a similar notion of neutralized grammar exists in a handful of nouns that refer to both sexes, which are used with a unique gender marked article (e.g., *une*_{feminine} *personne* [a person]), in terms of grammatical status, they are not truly nominalized forms and are not frequently found.

Considering the discrepancies between the languages, it is questionable as to what extent bilinguals are constrained by the representational restrictions present in each of their languages. Indeed studies have shown that the different structures of the bilingual's languages interact and interfere with one another. For instance, Sato, Gygax and Gabriel (2013) found that gender representation among bilinguals of French and English leaned toward the representational tendency of the language in use, although in the L2, the influence of their first language (L1) was nonetheless observed. In terms of selecting relevant linguistic systems, these results concur with theories suggesting that bilinguals unselectively activate both of their languages irrespective of the language being used for language processes (de Groot et al., 2000; Dijkstra & van Heuven, 1998). The comparison of these two language groups therefore provides a basis to further qualify the impact of gender-neutral language.

6.2.4 The present study

The present study aims to add to the sparse literature on the impact of gender-neutral language use on readers' mental representation of gender. We propose a study in order to examine the impact of language by further scrutinizing the stimulus material and participant sample. As nominalization forms in German are grammatically dissociated from gender, they should not activate a specific gender. Essentially, when reading German nominalization forms, we expect a reduction in the persisting male bias triggered from reading the masculine form in grammatical gender languages. For non-native German readers, a similar male-attenuation effect is also expected in their L2 German, which should provide more direct evidence of language as a common source for bringing about changes in gender biases. Our non-native German sample was comprised of French L1 readers whose L2 German was well established (see Results section for L2 proficiency). With French as their L1, these participants were familiar with the notion of the nominalized form, although they are gender specific in their L1 French. More importantly these participants were familiar with the concept of the generic interpretation of the masculine form. In Sato et al. (2013), L1 French readers transferred their language-bound male biases to a non-grammatical L2 (i.e., English). These results suggested that language is a strong source for the representational shift, although the magnitude of the transfer depended on L2 proficiency level, with proficient L2 speakers showing less impact of their L1. Based on these findings, we can assume that non-native German readers with lower proficiency may in fact maintain the default male biases associated with person reference role nouns inherent in their L1. For highly proficient readers, nominalized forms may show a neutralizing effect in gender biases.

The study differs from past research in the following three ways. First, we applied a more stringent definition of grammatically neutral forms by specifying nominalized forms. Past studies have applied related, yet different, assorted grammatical forms as a neutral grammar condition (e.g., Braun et al., 1998). These nominalized forms were then matched in semantic context with masculine forms to attain a better comparison of the two grammatical forms (e.g., masculine form: *Die Käufer* [the buyers]; nominalized form: *Die Konsumierenden* [those that consume]). Second, our empirical focus was purely grammatical, comparing nominalized with generic masculine forms, as opposed to the study by Irmen and Roßberg (2004), which also examined effects of gender stereotype as well as grammatical influences. Eliminating other gender-associated information from the experimental paradigm provides more convincing support for the debate on the promotion of gender-neutral language, as the heart of the discussion rests upon how linguistic (i.e., grammatical) reform may bring about a shift in people's mental representations. Third and finally, the present study adapted both objective (i.e., response times) and subjective (i.e., response proportions) measures to gauge gender-processing effects at different levels of processing. While the former is representative of less-monitored processing (i.e., automatic), the latter reflects readers' subjective judgments of gender information. A sentence evaluation paradigm that provides both these measures was thus implemented for the study.

6.3 Method

6.3.1 Participants

Forty-three native German-speaking students (mean age = 21.88 years, age range = 19 – 42, 2 males¹⁹) and 43 non-native German-speaking students (i.e., native French-speaking) (mean age = 21.47 years, age range = 19 – 40, mean start age of L2 acquisition: 9.04 years; range = 0 - 13 years; mean years of L2 study: 9 years; range: 6 -12 years, 7 males) participated in the experiment. All students were recruited from the University of Fribourg (Switzerland) and were granted course credit for participation.

¹⁹ As past studies assessing masculine plural forms as generic have never reported any participant sex differences in evaluating similar sentence paradigms, we did not control for their sex.

6.3.2 Design and Materials

6.3.2.1 Sentence evaluation task

The experimental task followed the sentence evaluation task initially proposed by Tanenhaus and Carlson (1990) and later adapted in numerous reading comprehension studies (e.g., Garnham et al., 2012; Garnham, Oakhill, & Cain, 1997; Garnham, Oakhill, & Reynolds, 2002; Kurtzman & MacDonald, 1993). In this task, participants read a passage, one sentence at a time, and evaluate whether the final target sentence is a good continuation of the previously presented sentences. Differences in judgments (i.e., yes / no responses), as well as response times, reflect the difficulty or ease in which readers map information onto their mental representation.

In our study, as in Garnham et al. (2012), participants were asked to read a composite of three sentences. The first sentence introduced a group of people denoted by a role noun, the second specified what they were doing, and the final sentence specified that there were “*some (of the) men*” or “*some (of the) women*” in the group referred to by the role noun in the first sentence (see sentences [1a] to [1c] in the example below). The manipulation of grammatical information was done by presenting the role nouns either in the masculine or in the nominalized plural form, both potentially interpretable as generic. It was presumed that positive *yes* responses to target sentences with female continuations (i.e., “*some (of the) women*”) were indicative of a more flexible interpretation (i.e., not a gender-specific interpretation).

As we aimed only to examine the effect of grammatical information (i.e., masculine vs. nominalized forms), only role nouns that did not entail a gender stereotype were chosen. Furthermore, provided that the derivations of the two grammatical forms differ, only pairs of masculine and nominalized role nouns that shared similar semantic context were selected to maintain compatibility. To adhere to such criteria, 20 participants (who did not participate in the main experiment) estimated, using an 11-point rating scale (as done in Gabriel, Gygax, Sarasin, Garnham, and Oakhill, 2008 and Misersky et al., 2013), the extent to which a list of 133 role nouns were made up of either men or women. One side of the scale corresponded to 100% male composition and the other side to 100% female composition. The scales were inversed for half of the participants (stereotypicality rating: $M = 5.63$, $SD = .5$). Of the role nouns that received a high neutral rating (i.e., 50% women), 32 role noun pairs (32 for each grammatical form, hence 64 in total) were made that matched in semantic relatedness. Twenty new participants assessed the semantic relatedness of each pair on a 7-point scale (1 = *highly*

unrelated to 7 = *highly related*). Finally, twenty stereotypically neutral pairs of role nouns (40 role nouns in total: see Table 6.1) that were judged as having the highest semantic relatedness scores (relatedness rating: $M = 5.32$, $SD = .9$) were chosen for the experimental stimuli. All masculine plural forms ended with *-er* suffixes (e.g., *die Käufer* [the buyers]) and the nominalized forms with *-en* suffixes (e.g., *die Konsumierenden* [those that consume]).

(1a) *Die Käufer / Die Konsumierenden* waren schon im Restaurant.

[*The buyers / The consumers* were already at the restaurant.]

(1b) Sie aßen.

[They ate.]

(1c) Es war offensichtlich, dass ein Teil der *Frauen / Männer* gut gelaunt war.

[It was obvious that some of the *women / men* were in a good mood.]

To ensure that participants did not see two role nouns with the same semantic context (i.e., two role nouns from a pair) within the experiment and the same role noun with both gender continuations, we made four lists of 20 experimental passages, each with ten masculine and ten nominalized forms. Each participant was randomly given one of the four lists. If a role noun appeared in one list, its associated role noun, in the other grammatical form, would appear in the other list, and if a role noun appeared in one list followed by a female continuation, it would appear in the second list followed by a male continuation. In order to prevent grammatical structures from encouraging and emphasizing certain biases, the first half of the experiment was presented in either the masculine or nominalized condition first. Afterwards, the grammatical condition switched to the other half. Hence, the order of the grammatical condition was counterbalanced with the remaining third and fourth lists.

Again, both grammatical types could be interpreted as generic, and therefore, participants could potentially answer *yes* to both gender continuations following both grammatical types. If participants responded *no* to a gender continuation, this essentially suggests difficulty when interpreting the role nouns as generic (i.e., inclusive of both genders). Finally, 20 filler sentences that elicited a clear *no* response were added to each grammatical form condition. The structure of the filler passages was identical to that of the experimental passages but established a semantic inconsistency (see example [2a] and [2b] for an English example). The fillers were identical for each list.

Table 6.1 Non-stereotypical role nouns chosen for experimental items

English Translation	Nominalized Forms	Masculine Plural Forms
The married / spouses	die Verheirateten	die Ehepartner
The onlookers	die Schaulustigen	die Gaffer
The employees	die Angestellten	die Arbeitnehmer
The teachers	die Lehrpersonen	die Lehrer
The art lovers	die Kunstbegeisterten	die Kunstliebhaber
The consumers / buyers	die Konsumierenden	die Käufer
The tobacco addicts / smokers	die Tabak-Süchtigen	die Raucher
The learners / students	die Lernenden	die Schüler
The laypeople / beginners		die Anfänger
The animal lovers	die Tierliebenden	die Tierliebhaber
The diligent	die Fleissigen	die Streber
The doctors / academics	die Promovierenden	die Akademiker
The runners	die Rennenden	die Läufer
The old / retired	die Alten	die Rentner
The foreigners / strangers	die Fremden	die Ausländer
The trainees / employees	die Auszubildenden	die Mitarbeiter
The responsables / managers	die Verantwortlichen	die Betreuer
The locals / residents	die Einheimischen	die Einwohner
The music-fans / musicians	die Musikbegeisterten	die Musiker
The travellers / frequent flyers	die Reisenden	die Vielflieger

(2a) Die Professoren verbrachten die Pause in der Sonne.

[The professors spent the break in the sun.]

(2b) Sie genossen das schöne Wetter.

[They enjoyed the nice weather.]

(2c) Wegen des schlechten Wetters hatte die Mehrheit der Frauen einen Regenschirm.

[Because of the bad weather the majority of women had an umbrella.]

Experimental and filler items were randomly presented. All items were presented in German and each participant saw only one list.

6.3.2.2 Language proficiency evaluations

To evaluate non-native German proficiency, a German C-test from onDaF (www.ondaf.de/) was used. The C-test is a modified form of a cloze test that presents four to five texts where the second half of every other word, apart from in the first and last sentences, is masked. Participants are required to fill in the blanks with the amount of correct restorations providing a measurement of their overall efficiency to process language (Grotjahn, Klein-Braley, & Raatz, 2002). In the C-test we employed, the participants' task was to reproduce 25 blank words in four texts in 20 minutes. Each participant was subsequently instructed to self-evaluate their proficiency levels for each language competence (i.e., listening, reading, speaking, writing) on the European Framework of Reference (Council of Europe, n.d.) on a questionnaire which also asked about their language background (e.g., age of acquisition).

6.3.3 Apparatus

Passages were presented using the Psyscope Software (Cohen, MacWhinney, Flatt, & Provost, 1993) on a Power Macintosh 4400 computer. A button box (with milliseconds accuracy) with buttons labelled “*Ja*” [yes] and “*Nein*” [no] was connected to collect participant response data. For each participant, the “*Ja*” button was adjusted so that it was always pressed by the participant's dominant hand.

6.3.4 Procedure

Participants were individually tested in a quiet room. Each passage began with a “***Bereit?***” [Ready?] prompt that prepared the participants for a new passage. Their task was to read the passages that appeared sentence by sentence on the computer screen in front

of them. Participants had to press the “*Ja*” [yes] button in order for the following sentence to appear after reading each sentence. The last target sentence was presented in blue print to visually indicate that participants had to respond by deciding whether this last sentence was a sensible continuation of the previously read sentences by pressing either the “*Ja*” or “*Nein*” button.

After completion of the main experimental task, non-native German speaking participants completed a C-test in German to scrutinize their proficiency in German. Additionally, they were given a list of role nouns that had appeared in the experiment to translate from German into French in order to identify any role nouns they were not familiar with. This was done in order to scrutinize any items that participants did not understand. Participants then completed a questionnaire to self-evaluate their German proficiency. Native German speaking participants did not complete the C-test and were only asked to identify role nouns they did not know in a given list. None were reported.

6.4 Results

Analyses were conducted only on the proportion of positive responses to the final sentence (i.e., “yes, this sentence is a sensible continuation”) and their response times²⁰. Essentially, all continuation sentences should have received positive responses, as these sentences were grammatically correct in relation to their preceding ones. However, when target sentences were difficult to process, participants responded negatively or with a longer processing time. For example, although the role noun in the first sentence could be interpreted as generic, a participant may have responded “no” (or “yes” more slowly) to a continuation including “some of the women” if they had interpreted a role noun as male-specific.

For non-native German speaking participants, items for which they did not provide a translation in the role noun check task were regarded as unknown and thus excluded from the analyses (in total, 19% of the data were excluded, with an average of six items per participant). We analyzed the data for each language group separately (as done in Sato et al., 2013) as the magnitude of the effects was expected to differ according to proficiency for the non-native group.

²⁰ Although being grammatical neutral, the item *Laien* (instead of *die Laienhaften*), sharing a similar ending as other nominalized items was mistakenly included as a nominalized form in the experimental design. The item was therefore excluded from the analyses.

6.4.1 Proportion of positive responses

The proportion of positive responses for each participant was examined with a mixed ANOVA on participants (F_1) and item (F_2) means. For the German native group, the by-participant analyses were conducted considering Experimental Order (masculine form presented before nominalized form vs. nominalized form presented before masculine form) as between-participants and Grammatical Gender (masculine form vs. nominalized form) as well as Gender Continuation (men vs. women) as within-participant variables. Item analyses were conducted with Experimental Order and Gender Continuation as within-item variables, and Grammatical Gender as between-item variables. Post-hoc tests were adjusted with Bonferroni corrections.

As language Proficiency was a vital variable among non-native speakers, it was also included as a between-participant variable and a within-item variable when analyzing data from the non-native group. To operationalize proficiency in the ANOVA analyses, the non-native German speaking group was divided into different proficiency groups (high vs. low), measured by the C-test and the self-assessment completed after the main experimental task. They were then split into two groups through hierarchical cluster analyses using Ward's method. This method results in a minimum increase in the error sum of squares within the generated clusters and hence enables a better method of grouping than the simple median split. The analyses produced a high ($n = 23$, $M = 82$, $SD = 9.18$) and a low ($n = 22$, $M = 57.32$, $SD = 9.86$) group that differed significantly ($W_s = 253$, $z = -5.75$, $p < .001$).

6.4.1.1 Native German speaking group

The analysis concurred with past studies showing a main effect for Gender Continuation, $F_1(1, 41) = 4.02$, $p = .05$; $F_2(1, 37) = 3.98$, $p = .06$, that denoted a greater proportion of positive responses for *men* (.78) than *women* (.72) continuations. Most importantly, the effect was qualified by a significant Grammatical Gender by Gender Continuation interaction ($F_1(1, 41) = 5.81$, $p < .05$; $F_2(1, 40) = 8.99$, $p < .01$) with post-hoc tests revealing that although participants manifested a greater male bias when role nouns were presented in the masculine form (men continuations = .80, women continuations = .66; $t(42) = 2.81$, $p < .01$) this bias disappeared when role nouns were presented in the nominalized forms (men continuations = .75, women continuations = .77; $t(42) = -.5$, *ns.*) (see Figure 6.1).

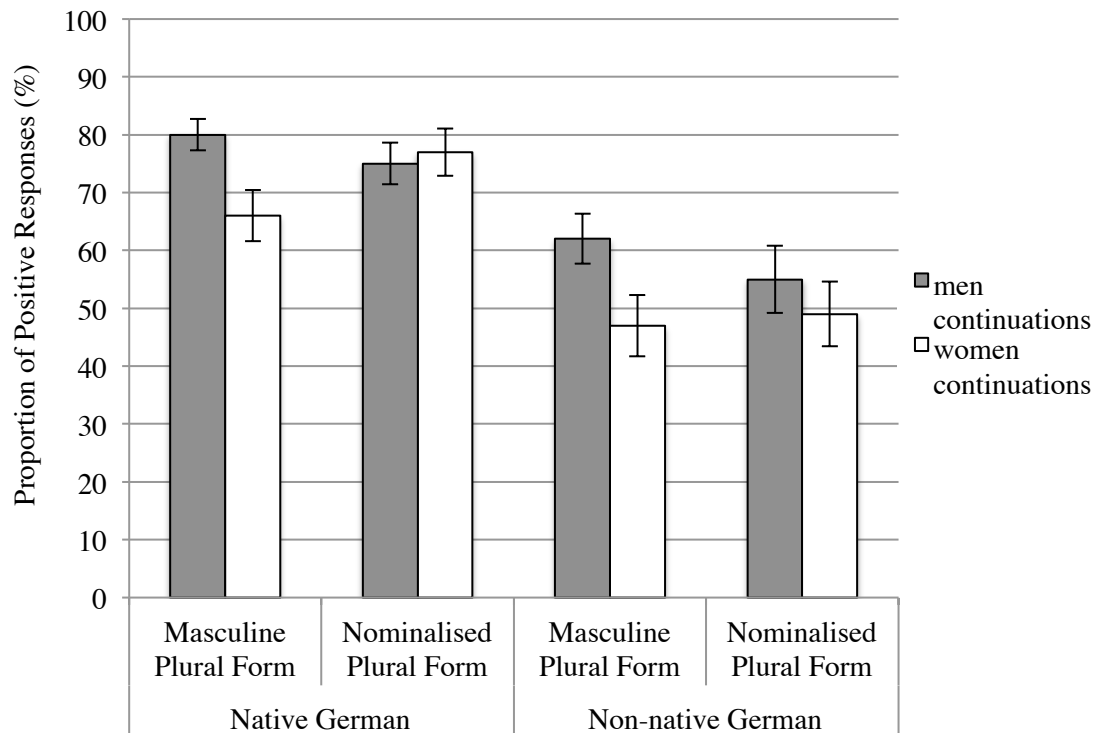


Figure 6.1 Proportion of positive responses for the native and the non-native German group in each grammatical gender form. Error bars indicate standard errors.

A significant Experimental Order by Grammatical Gender interaction surfaced, $F_1(1, 41) = 4.64, p < .05$; $F_2(1, 37) = 8.79, p < .01$. Post-hoc tests showed two interesting results. On the one hand, experimental order did not affect the proportion of positive responses for nominalized forms $t(41) = -.25, ns$. (presented before the masculine form: .77, presented after the masculine form: .76). On the other hand, masculine forms received greater positive responses when presented after (.81) than before (.66) nominalized forms, $t(41) = -2.79, p < .01$. Reading nominalized forms before masculine forms somehow influenced the acceptance of the latter, most likely augmenting their visibility in the disambiguation process. The effect however was not strong enough to completely alter the acceptance of specific gender continuations (as would be reflected in a higher order interaction including Gender Continuation). Given that the acceptability of masculine forms did not significantly increase when they were presented after nominalized forms, it would be unlikely that this effect was due to an experimental training effect.

6.4.1.2 Non-Native German speaking group

The analysis revealed a close to significant main effect of Gender Continuation, $F_1(1, 41) = 3.09, p = .09$; $F_2(1, 37) = 2.85, p = .10$, with participants showing a greater male bias

again, with more positive responses for men (.58) than women (.49) continuations, mirroring results obtained from the German native sample. The main effect of language Proficiency reached significance, $F_1(1, 41) = 9.01, p < .01$; $F_2(1, 37) = 21.62, p < .001$, with high proficient participants giving more “yes” answers (.63) than low proficient participants (.44). No effect including Grammatical Gender was significant (see Figure 6.1). More specifically, the Proficiency by Grammatical Gender by Gender Continuation interaction was not significant, $F_1(1, 41) < 1$; $F_2(1, 38) < 1$, suggesting that even with varying levels of L2 proficiency, grammatical form may not have substantial effects on non-native speakers to overturn male biases. This is in line with the Proficiency by Gender Continuation interaction, suggesting that regardless of the grammatical form, their male biases originate from their L1. No other main effects or interactions reached significance (all other $p > .1$).

6.4.2 Response times

Following common psycholinguistic practice, response times were fitted onto a regression equation for each participant with sentence length as independent variables (Trueswell, Tanenhaus, & Garnsey, 1994). This accounted for differences in individual reading speed, while also taking sentence length of the stimuli into account. Negative response times indicate a faster response time than predicted and positive response times indicate slower response times. Response times exceeding 2.5 standard deviations from each participant’s mean was excluded from the analyses (2% of data for the native German speaking group and 2% for the non-native German speaking group).

After the residual transformation, response time measures were analyzed using a linear mixed-effects model (Baayen, Davidson, & Bates, 2008) with R (R Core Team, 2013) and the *lme4* (Bates, Maechler, Bolker, & Walker, 2013) and *languageR* (Baayen, 2013) packages, mainly to account for unbalanced data (as only “yes” responses were considered). The four predictors used in the analyses of the proportion of positive responses were simultaneously entered as fixed effects (i.e., Gender Continuation, Grammatical Gender, Experimental Order and Proficiency for the non-native German speaking group). C-test scores, which indicate language proficiency, were centered and included as a continuous predictor. All other predictors were sum coded (1, -1).

Each predictor was added into the model by comparing the simpler model to the more complex model using log-likelihood ratio tests. By doing so, we removed any predictors or their interactions that did not improve the model (Baayen et al., 2008). Participants and items were entered as random effects and random slopes with a maximal random factor structure

(Barr, Levy, Scheepers, & Tily, 2013). *P*-values for fixed effects of the final model were obtained through Kenward-Rogers approximation and post-hoc tests for interactions were corrected with Bonferroni adjustments with the *multcomp* package (Hothorn, Bretz, & Westfall, 2008).

6.4.2.1 Native German speaking group

The final model included the main effects of Gender Continuation, Grammatical Gender, Experimental Order and the Grammatical Gender by Experimental Order interaction and by-participant random slopes for the effect of Experimental Order.

A significant effect of Gender Continuation revealed that, as expected, responses to men continuations (-175 ms) were faster than women continuations (48 ms), $F(1, 589.17) = 5.8, p < .05$). Although the main effect of Experimental Order was not significant ($F(1, 38.57) = 2.55, ns.$), the main effect of Grammatical Gender was marginally significant, showing that participants responded faster to continuations following nominalized forms (-127 ms) than masculine forms (19 ms) $F(1, 589.60) = 3.16, p = .08$. This effect was further qualified by a significant Grammatical Gender by Experimental Order interaction $F(1, 578.93) = 12.31, p < .001$. When the masculine form was presented before the nominalized form, acceptance of continuations that followed masculine forms were slower than continuations that followed nominalized forms (masculine forms: 90 ms, nominalized forms: -381 ms, $p < .05$). However, responses to both grammatical forms did not differ when the nominalized form block was presented before the masculine form block. In other words, the continuations preceded by nominalized forms were responded to faster, irrelevant of the nature of the continuation (i.e., men or women continuations), when participants saw the masculine form first. The fact that this effect was present for both men and women continuations suggests that reading a certain grammatical gender form *first* (i.e., beforehand) simply *generally* facilitates disambiguation of the grammatical form presented second (i.e., less processing time needed). It does not however facilitate the activation of a particular representation. The Grammatical Gender by Gender Continuation interaction found in the response proportions did not significantly improve the model fit and was therefore not included in the final model.

6.4.2.2 Non-native German speaking group

Following the native German speaking group, the final model for the non-native group also included Gender Continuation, Grammatical Gender and the Grammatical Gender by Experimental Order interaction. The random structure included random intercepts for

participants and items as well as by-participant random slopes for Grammatical Gender. The Proficiency predictor did not elicit a significant main effect or interact with any other predictors. No other main effects or interactions were significant.

As suggested by results of the native German group, the Gender Continuation showed a considerable effect indicating that men continuations (-105 ms) were again positively responded to considerably faster than women continuations (186 ms), suggesting robust male-biases $F(1, 285.69) = 5.74, p < .05$. The effects of Grammatical Gender $F(1, 40.54) = .42, ns.$, and Experimental Order $F(1, 35.61) = .14, ns.$, were not significant. However, the interaction between Grammatical Gender and Experimental Order was highly significant, $F(1, 38.51) = 10.3, p < .01$. Participants sped up their responses to grammatical forms that were presented afterwards compared to those that were presented beforehand ($p < .05$). Although this interaction effect appears similar to that found for native German speakers, we believe that these effects are indicative of a training effect. For non-native German speakers, the grammatical gender form presented in the second half of the experiment was always responded to significantly faster than the grammatical gender form presented in the first half of the experiment. For native German speakers, this was not the case. Responses were not faster when masculine forms were presented after the nominalized forms.

6.5 Discussion and conclusion

Recent research on gender representation has provided consistent evidence that in grammatical gender languages such as German, mapping the male gender to a role noun is facilitated by the use of the default masculine form, whether it is intended as generic or not. This male bias suggests that information embedded in the surface structure of a noun has a definite role during gender representation. In the present study, we aimed to assess whether such a robust bias could be alleviated by employing neutral grammatical forms such as nominalized forms of adjectives and participles (i.e., nouns that are derived from adjectives and participles), and whether the main reliance of surface information still holds for non-native speakers.

The results of our study indicated a predicted and consistent global male bias, which was observed for both native German and non-native German speakers. This effect, however, was reduced when considering positive response proportions for native speakers when role nouns were presented in the nominalized forms. These findings are consistent with our hypothesis that surface information, such as grammatical markings, has a substantial impact

on readers' decisions about gender: grammatical masculine forms evoke male-biased representations and grammatical neutral forms evoke gender-neutral representations. These results speak to the idea that the language in which information is encoded guides readers' attention to specific pieces of information, which in turn grounds the focus of comprehension processes.

For native German speakers whose grammatical gender system is well-established, the grammatical form that carries a gender-neutral association appears to exert an influence that changes mental representations of gender. However, contrary to our hypothesis, this interaction was not found for non-native speakers who showed a persistent male bias even when role nouns were presented in the nominalized forms in their L2. As the analyses were only conducted for items that participants understood, this finding cannot be explained by a lack of understanding of the role nouns. For non-native speakers this particular grammatical form does not seem to exert a sufficiently sizeable effect to change mental representations of gender. Consequently, male biases remain. These results may be attributed to the general lack of use of gender-neutral forms (as opposed to more commonly used gender-fair forms such as pair forms) in society. These findings concur with those of Sato et al. (2013) who reported that less-proficient L1 French speakers transferred the French-associated male bias to L2 English when processing role nouns in English (i.e., yet with no masculine marking).

With regard to theories of bilingualism, these results also support the fact that both the L1 and L2 were highly activated irrespective of whether the experimental task was conducted in the L2 (de Groot et al., 2000; Dijkstra & van Heuven, 1998). Male biases associated with the default grammatical gender thus appeared to provide the general representational gender bias even when processing role nouns that lacked a specific grammatical gender in their L2. While the data for the non-native German group did not provide the expected results to support the thinking-for-speaking effects in L2, these effects remain interesting. Despite the understanding of the general German grammatical gender system (as present in their L1), as well as the specific presented nominalized stimuli, non-native German speakers were not influenced by surface-level information. These results suggest that L2 grammatical categories that differ in their function to those in L1 are more difficult to integrate and to eventually represent, particularly in cases where the interpretations between the two languages differ. Furthermore, the contribution of surface forms of specific grammatical categories in terms of representations is much more reduced than it is for native speakers.

An important point that should be noted here, especially among native German speakers, is the inconsistency of the results surfacing in our two measurements. The impact of

the nominalized forms, as seen in the significant grammatical gender by continuation interaction, was reflected among German native speakers in their response proportions, yet not in their response times. As was mentioned earlier, response proportions were considered to reflect readers' subjective judgments about gender whereas response times were expected to reflect less-controlled processes. The fact that the response time measurements only indicated a male bias and not an interaction effect with grammatical form could suggest the absence of an automatic process to map neutral, non-gender representations. Although our participants were fully acquainted with the nominalized forms tested, a few more years of exposure to these forms may be required for their processing to become automatic.

Nonetheless, these lines of evidence lend strong support for the idea that language is deeply responsible for guiding readers' attention to specific pieces of information (during native language processing) given that readers need to attend to these linguistic regularities of the language. While the activation of a specific gender may not essentially be necessary to maintain local text coherence, the representation of gender has been found to be automatic and difficult to avoid (Banaji & Hardin, 1996; Oakhill et al., 2005; Reynolds, Garnham, Oakhill, 2006). However, the results of the present study suggest that although nominalization forms do contribute in suppressing gender biases, the maintenance of such gender-neutral representations are difficult, particularly for non-native speakers.

Our results also showed a consistent Experimental Order by Grammatical Gender interaction for both language groups. For native German speakers in particular, response times showed that responding to the masculine form was not influenced by presentation order as readers are always faced with the task of disambiguating its interpretation (i.e., male specific vs. generic interpretation). Resolving other competing information may hence lead to difficulty irrespective of experimental order. In contrast, nominalization forms were responded to faster when the masculine form was presented first, suggesting that reading a certain grammatical gender form *first* (i.e., in the first block) *generally* facilitated disambiguation of the grammatical form presented second (i.e., less processing time needed). Additionally, the low positive response proportion for masculine forms that were presented first further confirmed the general representation difficulty associated with disambiguation

With regard to the response proportions, although responses to nominalizations did not differ in function of experimental order, masculine forms did (i.e., increased acceptance when presented after nominalized forms). Reading nominalized forms could have triggered participants' sensitivity to the generic interpretation, which may have persisted when participants switched to masculine forms. Consequently, both specific *and* generic

interpretations might have been equally activated when masculine forms followed the nominalized forms. In fact, this is only partly supported by the data, as both men and women continuations received more positive responses after masculine forms presented in the second part. However, the male bias was still apparent when masculine forms were presented after nominalized forms, as signaled by a non-significant higher-order interaction between Experimental Order, Grammatical Gender and Gender Continuation.

Processing grammatical gender of person nouns is intricate, as information associated with its surface structure influences gender representations. The current study showed that the use of nominalized forms can alleviate the robust male bias in grammatical languages, at least when *more* controlled processes are concerned. Although we observed a limited impact of these linguistic forms on less competent non-native speakers, they provide an attractive substitute for gender-balanced language forms when possible. Furthermore, we believe that assessing new gender-fair forms is crucial in an era where women's visible participation in society is fundamental.

Chapter 7

General Discussion

7.1 Introduction

The process of reading comprehension imposes several challenges on readers, especially during the on-line construction of mental representations. One critical issue consists of how specific languages influence their readers to focus on certain aspects of the language, given that the regularities of how each language encodes specific facets of information may require readers' heightened attention to the information. Attending to these specific regularities may consequently influence how readers represent the information during the construction of their mental models (Slobin, 1996a, 1996b, 2003). Studies employing bilinguals who communicate in two languages have shed light on the fact that when the bilinguals' two languages deviate in the way they lexicalize specific information, their understanding in each language may very well reflect such language specific features (e.g., Athanasopoulos & Kasai, 2008; Athanasopoulos, 2006, 2007; Berman & Slobin, 1994). This issue was further explored in the present thesis to provide a better understanding of how bilinguals switch their representational tendencies according to either of their languages. In particular, the focus was centered on how bilinguals make use of gender information pertaining to their world knowledge (i.e., gender stereotypical knowledge) and grammatical information (i.e., grammatical gender).

The reported studies have worked under the basic assumptions of research on gender representation offered by Gygax et al. (2008) that revealed processing tendencies in English, French and German readers. Their central findings illustrated that gender stereotype information played a key role in representations for English readers, and that the grammatical gender marking available to French and German readers had an overruling effect on processing mechanisms. Importantly, although masculine forms are generally used as a default gender marking to refer to both male-specific and generic connotations, the representations were found to be male-biased.

Based on these assumptions, the present thesis examined whether these divergent representation tendencies in each of the bilinguals' languages would impose a modification in gender representation as a function of language shift. Importantly, bilinguals' L2 proficiency

was expected to play a guiding role in the shifting process in the L2, with the processing tendency leaning towards native-like representations as their proficiency increased. Ultimately, the studies aimed at providing a better understanding of the relationship of language and cognitive processes, specifically in how social perceptive biases are activated and altered as a function of shifting languages. The present chapter revisits the main findings that emerged from the experimental chapters, extending the discussion on research implications and finally addressing the limitations and directions for future research.

7.2 Summary of the main findings

Study I examined whether the specific characteristics in conveying gender information in English and French would result in different mental representations among French-English and English-French bilinguals. A similar sentence evaluation task employed in Gygax et al. (2008) was adapted to address both the bilingual's L1 and L2. The results suggested that readers largely conformed to specific language-bound tendencies for gender representation in each language. A switch to English resulted in the activation of gender stereotypical information which guided the means to represent gender information. On the other hand, a switch to French led to a greater impact of the information conveyed by the masculine grammatical marking than stereotype information, with readers activating male-biased inferences to their representations. As representations altered according to how readers shifted between their languages, the study provided strong evidence to suggest that language played a vital role in influencing cognitive processes.

Importantly, processing discrepancies were seen *within* the L1 French group that concurred past studies on bilingual categorization (Athanasopoulos & Kasai, 2008; Athanasopoulos, 2009; Phillips & Boroditsky, 2003), wherein L2 proficiency was found to be a good indicator to predict the magnitude for which the male biases emerged. Higher L2 proficiency in English suggested that readers based their gender representation more on English relying on stereotype information, whereas lower L2 English proficiency indicated a greater male bias on the representation carried from their L1 French to the L2 English. In contrast, the native English group did not demonstrate such proficiency-dependent differences given that the concept of grammatical gender did not exist in the L1 that rendered the internalization of the grammatical gender system difficult for readers with a non-grammatical gender language as an L1.

The findings from Study I revealed that during reading processes, bilinguals of natural gender and grammatical gender languages activated gender information based on varying sources of information in each of their languages. Specifically for the latter, surface level associations via the masculine forms appeared to serve as male-emphasizing cues that impacted the resulting representations. In Study II, the influence of these surface level cues was further investigated in more detail using a linguistic-visual paradigm. Consisting of a single-word prime that was immediately followed by a visual judgment task, this paradigm was empirically and cognitively more ecological than linguistic paradigms. Not only did it allow an examination of the immediate effects of gender activation on a lexical level, it eliminated possible impacts of language processing resulting from the presentation of verbal target stimuli.

Particularly in Study II, we aimed to reduce the male-biasing effects observed in the masculine forms of Study I by presenting readers with other biasing surface forms. Specifically, the German plural determiner (*die*) that was expected to have female-associated biases was compared to that of the French plural determiner (*les*) that did not share similar surface associations. The findings confirmed the immediacy of both gender stereotype and grammatical gender information activation that were triggered even when discourse information was not provided. Unlike Gygax et al. (2008) who did not find significant stereotype effects in the processing of stereotypically biased role nouns presented in the masculine form, the effect was significant in Study II, most likely emerging from the lexical-based paradigm that was able to capture the immediate effects of stereotypes. However the impact of the German determiner that was expected to moderate the male-dominant representations did not emerge in either language group, further confirming the vigor of the activated male biases.

Importantly, the results revealed that processing in one's L1 triggered a greater male bias than in L2. In parallel to these results, higher levels of L2 proficiency as evaluated by C-tests were linked to higher competence in using the generic usage of the masculine form in one's L1. In contrast, lower levels in L2 were associated to inability to override the L1 male biases when processing in the L2. These results were important given that both of the tested languages were grammatical gender languages with a potential to carry male biases. This male bias originating from the L1 suggested that the gender biases triggered during gender representation are determined by readers' ability to control the languages in which the information is encoded.

The findings from Study II established that role nouns acted as strong cues activating both grammatical and stereotypical gender information. These findings led to Study III which addressed whether a specific gender would still be established when reading less salient gender cues that do not contain whole gender features (i.e., gender cues that carry only particular dimensions of person concepts). In this study, occupational descriptions associated to stereotypically biased role nouns were employed in two experiments employing an eye-tracking paradigm to test French-English bilinguals. The findings from the L1 data confirmed the strong gender effects that stereotypical role nouns activate. L1 French data suggested that both stereotypical and grammatical information were activated that generated gender inferences that were made elaborately when reading role nouns. In contrast, occupational descriptions *alone* were not strong enough cues to immediately activate gender, although when the text required the reader to search for a person reference by means of pronoun resolution, the occupational descriptions were nonetheless found to be associated to a specific stereotypical gender.

Contrary to expectations, findings in the L2 English did not reveal shifts in representations as Studies I and II did. Although male occupational descriptions compared to other stereotypes facilitated information integration processes, the overall anaphor resolution task appeared to be difficult in the L2. These effects may be explained by the fact that general L2 competencies of our participant sample were insufficient and processing for high-level information (i.e., anaphora resolution) was not properly accounted for. However, high proficient readers appeared to have taken more effort to represent gender information as reflected in the overall reading times that were longer than those of less-proficient participants.

Study III revealed that again, gender stereotypical role nouns presented in the masculine form in L1 French activated male-specific representations, and furthermore, despite the fact that gender information was not automatically elicited, even less salient cues were associated to readers' representation for a specific gender category. In Study IV, an attempt to temper the reported recurrent male biases was made, by testing the nominalization forms in German. French-German bilinguals (non-native German speakers) were compared to native German readers partly because no similar grammatical form exists in the French language and also because they allowed an examination of whether these cases could be newly and easily learnt. The findings again reflected that habitual usage of a specific linguistic form resulted in encouraging comprehenders to focus on specific characteristics emphasized by the language.

For L1 German speakers, the usage of the nominalizations proved to be an effective reduction measure to weaken the male bias as readers were linguistically obligated to deter from any specific gender biases. In contrast, while the French-German bilinguals indicated that they understood the nominalization forms in their L2 German, the lack of their habitual usage of these forms lessened the neutral impact the linguistic form conveyed. Importantly, a male bias persisted, which was found to be originating from their L1 French. Consequently, reading the nominalization forms in the L2 was not a sufficient means to alter the activated male bias (of the L1 French) irrespective of L2 proficiency levels. These results indicated the difficulty of maintaining a neutral representation as well as learning new linguistic features (nominalization forms) that do not exist in the readers' L1.

In sum, the findings provided a strong case to suggest that languages have a strong impact in influencing the mental representation of gender that may result in different representational outcomes according to the languages bilinguals employ for comprehension processes. The impact of language-bound regularities in readers' L2 was found to largely depend on readers' L2 proficiency that acted as an indicator to suggest how well the reader may control the recurring biases inherent in their L2. However when the L1 lacked the linguistic features and forms present in the L2, this effect did not show a persistent effect (e.g., Study I: English lacking grammatical gender; Study IV: French lacking the same functions of German nominalizations).

7.3 Theoretical and applied implications

7.3.1 Implication of findings for L1 gender representation

The findings of the studies presented in this thesis largely contribute to the notion that gender biases that emerge in readers' representations can be attributed to the different means languages encode gender. The different representation outcomes largely conformed to the language-specific tendencies outlined by Gygax et al. (2008) that indicated a strong dependency on stereotypical information in English, as opposed to a greater reliance of grammatical information in French and German.

In fact, the activation of gender information appeared to be situated in role nouns themselves, where an associated gender and its descriptions were contained as part of the lexical representation. Reading them essentially triggered an activation of a specific and associated gender that was automatically integrated into the mental model. If information

following the role noun suggested a gender that was inconsistent with the mentally represented gender, readers underwent an updating of their representations, resulting in a gender mismatch effect as observed in Studies I and III.

These lines of evidence are consistent with the mental models view proposed by Garnham (2001) that demonstrates that information about a person's gender is activated through an inference generation mechanism which is incorporated into readers' mental representations. Particularly, the mismatch effects seen in the dependent measures indicated that readers had already activated a particular gender prior to the textual specification of gender (as in Carreiras et al., 1996; Esaulova et al., 2013; Irmen, 2007; Reynolds, Garnham, & Oakhill, 2006).

Crucially, this activation mechanism was substantiated as occurring immediately when reading the role noun (Study II). Significant stereotype effects were observed even when no textual information supported the representation process. Such line of evidence lend support to psycholinguistic and social psychological studies illustrating that an activation of a social category automatically and immediately activates associated stereotypes (e.g., Dovidio et al., 1986; Hamilton & Sherman, 1994).

Importantly in Study II, when the role noun was marked for grammatical gender, it resulted in an interaction between grammatical gender and stereotype information. Specifically if the gender of grammatical and stereotype information matched, that is, when the masculine grammatical marking was accompanied by male stereotypical information, the representation of the male gender was considerably facilitated (Studies II and III) whereas a combination of the female stereotype with the masculine marking was found to hinder the representation process.

However, as reflected by the non-significant Stereotype effect in Study IV (as well as Gygax et al., 2008), the stereotype influence in grammatical gender languages was not as extensive to overcome grammatical gender information, and its effects appeared to have dwindled rapidly into the discourse context, that may have reduced its statistical power. In fact, this may reflect the reason why a significant stereotype effect emerged only when using a lexical based paradigm (Study II) where the immediate effects could be captured, as well as in an eye-tracking paradigm (Study III) where the measures were more sensitive to the fine-grained temporal aspects of on-line comprehension.

In contrast to the short-lived stereotype effects in grammatical gender languages, all studies (Studies I - IV) showed that in the presence of grammatical gender information,

gender interpretation favored the information conveyed by grammar over stereotypical information. Note that this did not necessarily result in a generic representation, but more in a male-dominant one, provided that the surface forms of the masculine form boosted the association to the male gender as has been reported in past studies (Braun et al., 2005; Gabriel & Gygax, 2008; Gygax et al., 2008; Lévy et al., 2014; Nissen, 2013). It appears that the masculine marking of the role nouns semantically contributed to a representation of men, on the grounds that on a surface level, the masculine grammatical form emphasized the association to the male gender. These findings corroborate findings reported in gender categorization studies that have shown that the knowledge of a grammatical gender system does not simply play a categorical function, but may heighten associations to conceptual gender (e.g., Boroditsky et al., 2003; Sera, et al., 1994).

Importantly, this male-bias was found to be robust, being resistant to probable female biases that were expected to emerge from associated surface forms of German plural determiners (Study II). In fact, a male-dominant representation could only be attenuated when the masculine form was replaced by non-masculine (and truly generic) forms such as nominalization forms (Study IV), which explicitly hindered readers' attention from activating any information related to gender.

All together, the findings corroborate the thinking-for-speaking hypothesis (Slobin, 1996a, 1996b, 2003) that affirms that language plays a role in emphasizing specific attributes of reality. According to this view, if specific languages impose a *habitual* encoding of a specific domain such as gender, readers will be encouraged to direct their attention to these aspects of the language. Reading in a grammatical gender language will in essence require the readers to assess this aspect of information the language emphasizes. Consequently reading in a language such as English where such requirements are not needed would result in readers relying on non-linguistic knowledge, in other words, world knowledge for gender information. However inasmuch as grammatical gender languages impose a male bias, readers' world knowledge also carries gender biases that become reflected in their representations. In this respect, the languages bilinguals speak have direct consequences on how they represent gender. Both the knowledge of grammatical and stereotypical gender appears to facilitate the representation of a specific gender included within the lexical representation of a role noun, with the former grammatical knowledge imposing a dominant impact on the latter stereotypical knowledge when both information sources are available.

7.3.2 Implication of findings for L2 gender representation

Results obtained from the presented studies largely indicated that a switch to the L2 language brought a shift to the gender representation tendency for bilingual readers. Particularly, the representation shift substantially surfaced in Study I employing French-English and English-French bilinguals, most likely due to the greater typological differences in which English and French encoded gender compared to the relative difference existent between the two grammatical gender languages (i.e., French and German).

The observed representational shift justified the thinking-for-speaking hypothesis, which postulates that comprehending in a specific language brings readers to attend to the specific regularities languages impose. Reading in French essentially required a greater perceptive awareness to attend to grammatical gender information, augmenting readers' predispositions to gender characteristics. Consequently, this awareness resulted in a facilitation mechanism to specifically represent the male gender in French, although such perceptive attention was not required in English, giving rise to diverging representations in each language.

These findings concerning representation shifts corresponded with the findings in bilingual studies validating the effects on categorization switch according to the language at use (e.g., Athanasopoulos et al., 2011; Athanasopoulos, 2009). In their studies, a difference in color naming pattern between two languages led bilinguals to differentiate color according to the linguistic patterns, which in effect resulted in perceptive differences of the color spectrum. Their experimental tasks using color could be argued as being more perceptual in nature than the studies presented here, which allowed them to address more conceptual levels of cognitive representation. Notwithstanding these differences in experimental paradigms, the findings of these studies clearly suggest that linguistic processing mechanisms, which bring attention to real-life issues (i.e., the categorization of the female and male gender), were flexible and subject to change with language.

Crucially, the shifts to the target representation tendency observed in the studies were not complete, in other words, they did not entirely resemble the native group of the target language. Such effects were similar to the results reported in Athanasopoulos et al. (2011) who found different magnitudes of categorization tendencies shift in accordance to the participants' L2 proficiency and length of stay in the L2 speaking country. In the studies reported in this thesis, the magnitude in which the representation biases shifted was also found to be influenced by readers' L2 proficiency (Studies I and II). Higher fluency in L2

indicated that the shift would be greater than a representation shift in lower proficiency. The fact that such a gradual shift dependent on L2 proficiency was found suggested that these processing biases and tendencies were not bound to specific processing languages in an all-or-nothing manner but reveal that their impact changed according to the influence the language permeates to the reader. The transference of the representation tendencies according to language provided strong evidence to assert that language itself was a strong factor that shaped the readers' language comprehension mechanisms. In terms of language processing, a greater reliance on the L1 for low proficient bilinguals suggested that their L1 was more accessible to them during their L2 comprehension. If difficulties in comprehension processes were encountered, comprehension strategies that were stronger could have been employed from their L1 and applied to their weaker L2. Moreover, when linguistic features and functions present in the L2 did not exist in readers' L1, they struggled to implement these regularities when processing in the L2 as they were less accessible and acquirable (English native group in Study I, non-native German group in Study IV). These findings suggested that acquisition of new linguistic features and rules are difficult and that their impact may be reduced, allowing a greater impact of bilinguals' dominant L1 to come into effect.

This influence of L2 proficiency over representation shift was particularly interesting between the two grammatical gender languages given that both languages were expected to show a male bias. The findings from Studies II and IV which compared French and German revealed that the male bias triggered from the masculine grammatical form was stronger and more likely to manifest itself when reading in one's L1 than L2, with lower proficiency in the L2 showing an elevated impact of the male bias. The stronger male bias for low proficient participants was accredited to (i) their greater reliance for their male biased L1 than to the L2 as well as (ii) their lower competency to interpret the generic masculine as measured by an L2 C-test. To an extent, this was unexpected given that the L2 grammatical gender system was not new to these readers, but in fact were comparable to the grammatical gender system in their L1. More importantly, given that past studies in L2 reading have attested that there is a greater reliance for bottom-up processing than top-down processing, surface levels cues (i.e., masculine forms) that boost the male bias could have been expected to emerge in the L2. The results can be interpreted as demonstrating the relative complexity L2 comprehension imposed on readers. In fact, attention that was directed to representing the mentioned person reference might have already been a complex task for L2 readers. In this regard, the additional effort to attend to the new grammatical gender system could have been reduced.

Consequently, the impact of surface associations that appeared to have increased the male-bias in the L1 could have been reduced in the L2.

However, these results concur models of bilingual lexical access as illustrated by the Revised Hierarchical Model (Kroll & Stewart, 1994; Kroll & Tokowicz, 2005) which suggest that when L2 competency is still developing, the reliance to one's L1 is greater. On the other hand, when L2 is more developed, bilinguals develop greater autonomy from their L1. This would explain the greater impact of L1 seen among low proficient participants compared to those with higher proficiency levels.

These effects clearly highlight that for bilinguals, reading in the L2 had repercussions that impacted gender biases, which largely fluctuated according to their proficiency. The gender switch revealed from our data suggested that simple knowledge for certain words and sentences did not suffice in bringing about a shift in biases reflected in their representations. As suggested by the thinking-for-speaking hypothesis, the habitual usage and comprehension of the language was vital in bringing about a change in mental representations.

7.4 Limitations and directions for future research

The findings and discussions presented so far provide a foundation for better understanding the relationship between language and gender information. In the following, new directions for future research are addressed, elucidating the possible issues that were not covered in the present thesis.

7.4.1 Means to address linguistic relativity

A primary focus of this thesis was to examine whether differences in how gender information conveyed in languages could fundamentally influence how bilinguals represent gender, and whether these differences would ultimately influence gender perceptive biases. These aims were pursued based on the principle that gender information embedded in linguistic labeling of role nouns brings readers' attention to these aspects during the representation process. The outcome of this representation process has been linked to tangible consequences that may contribute to applied issues such as gender inequality (Prewitt-Freilino, Caswell, & Laakso, 2012), different degrees in reported job efficacy (Chatard et al., 2005) and children's vocational aspirations (Vervecken et al., 2013).

7.4.1.1 *Dissociating language from the cognitive processes in question*

To pursue these lines of questions, the studies reported in this thesis presented readers with psycholinguistic tasks that required comprehension of linguistic stimuli, subsequently followed by a judgment task or a comprehension question that allowed to capture the gender representation processes involved. With the core approach being centered on linguistic tasks (aside from Study II that employed a combined linguistic and visual conceptual approach), the research aimed at primarily addressing Slobin's thinking-for-speaking effects. This notion, also acknowledged in Gennari et al. (2002) as the language-as-strategy view, suggests that people apply language as an approach to engage and facilitate higher cognitive processes.

While it is claimed in this thesis that these influences of language deserve attention for further research, it acknowledged that they did not provide evidence to support the strong linguistic relativity effects. As Slobin (2003, p. 3) asserted, "Our basic cognition of gender does not change when we switch languages [...]. Although our social and cultural cognition may well change." Future research may hence approach the issue by shedding light on the direct impacts of languages on our non-linguistic cognition about gender.

As was delineated in Chapter 2, the biggest challenge for testing relativity effects concerned dissociating the effects of language from basic cognitive processing. As most basic cognitive mechanisms require language as a means to process information, the actual perception of an entity that may have been shaped by language is hard to separate from the surface level effects of language. A possibility of avoiding such confounds is to separate *linguistic* processing from the basic mechanism under question. Needless to say, this would involve non-linguistic stimuli such as visual imagery to be used as an experimental manipulation rather than the presentation of linguistic stimuli as done in the studies presented here. Additionally, participants would need to engage in a non-linguistic task (e.g., memory tasks, categorization tasks, etc.) and / or a linguistic intrusion task that would obstruct any (in)conscious linguistic processing to occur during processing. Studies employing both these measures examining how linguistic expression influences cognition have shown mixed results, confirming the need to further these experimental findings. For example, as previously described, Gennari et al. (2002) tested how differences of motion events in English and Spanish would influence non-linguistic tasks. No differences were found for recognition memory between conditions where participants watched video clips about motion events while repeating nonsense syllables (i.e., linguistic intrusion) and those who were instructed to describe the video clips. Nonetheless an effect on similarity judgment task was found in

conditions where participants were encouraged to linguistically encode the events. The authors took the evidence as reflecting the impact of language only when language was actively used for processing mechanisms, rejecting relativity theories. Contrary to their findings, Phillips and Boroditsky (Experiment 3, 2003) provided evidence to suggest that even under linguistic intrusion, grammatical gender effects were found to influence object perception. Bilinguals rated presented image pairs of persons (either female or male) and objects as being similar if the grammatical gender in the L1 and gender of the presented person image corresponded. The same findings emerged even when they were instructed to repeat random English letters that were played on an auto-stream that hindered participants from sub-vocally naming the objects.

7.4.1.2 Adapting experimental measures which are sensitive to the cognitive processes being tested

In order to increase the sensitivity of the processes being tested on-line, studies should also attempt to adapt different experimental measures. For example, Papafragou et al. (2008) tracked eye-movements while participants watched motion events with and without verbal encoding. They found that during linguistic encoding, events that could be linguistically encoded in the L1 were fixated more, although in conditions where verbal encoding was not required, no differences between the languages were found.

Thierry, Athanasopoulos, Wiggett, Dering and Kuipers (2009) were the first to examine ERPs in order to examine how languages impact visual streaming of color. They found that while English participants perceived various blue and green contrasts as being similar which were consistent with their linguistic encoding pattern, Greek participants who linguistically distinguish light and dark blue perceived blue contrasts different from green contrasts. They found that language differences occurred early and were considered as being partially perceptual as reflected by a Visual Mismatch Negativity (vMMN).

As seen in these examples, a focus on experimental manipulations taking the language variable into account, as well as on paradigms that are sensitive to temporal issues may reveal a better picture of the underlying processing of the influences of language on higher cognitive processes.

7.4.2 Investigation of individual differences

While the studies presented in this thesis concentrated particularly on the impact of specific linguistic cues in relation to participants' L2 proficiency, no formal attention was directed to other individual differences.

7.4.2.1 Individual stereotypicality

One critical variable that was omitted from the experiment was factoring in participant sex. Although attention to participant sex could be considered as being relevant given that the research was associated to gender cues, this choice to dismiss this variable was substantiated by past research in on-line gender representation that found no significant impact of the sex of respondents on the investigated processes. Besides the sex of respondents, there is ample evidence from social psychological studies that have indicated a close relationship between individual differences of held stereotypes and people's execution or attention to their prejudiced attitudes. Future research may examine the relationship between gender representation and linguistic cues in consideration to such individual factors.

For instance, Swim, Mallett and Stangor (2004) found that people's level of endorsement to Modern Sexist beliefs (as assessed by the Modern Sexism Scale) was associated to their inability to detect sexist language and more importantly, to their frequency in using sexist language. In light of these findings, the comprehension of stereotypically biased role nouns in grammatical gender languages which was one of the issues examined in this thesis, could be closely associated to participants' sensitivity to gender-fair language and their ability to detect that masculine forms also consist of a generic interpretation.

Similarly, Carter, Hall, Carney, and Rosip (2006) developed a measure capturing individual stereotype acceptance and demonstrated that individual acceptance to social stereotypes was found to correlate with different correlates that predicted a person's willingness to use their stereotypically held knowledge and beliefs for social interaction. Such measurements may account and illustrate a refined picture for the magnitude and differences within groups as to the extent participants activate stereotype information while processing gender stereotyped role nouns.

While participant differences in held stereotypes have been tested vigorously in off-line studies contributing to the understanding of social stereotyping, few have matched the results of questionnaires with on-line studies (e.g., Reali et al., 2015). Future research may

complement findings by accounting for individual differences to further assess the variance that may contribute to the linguistic impact of gender representation.

7.4.2.2 Individual cognitive load

Researchers generally agree that the activation of stereotype information facilitates person perception by increasing predictability of incoming person information. While research has provided evidence to suggest that these stereotype activation mechanisms occur effortlessly and automatically (Bargh, Chen, & Burrows, 1996; Fiske, 1998), the process involved in countering these stereotypes have been argued as being resourceful. Indeed studies have shown that when participants were given tasks that imposed a certain level of cognitive load, inhibition of stereotype activation was found to be difficult (Macrae, Bodenhausen, Milne, & Ford, 1997; Wyer, Sherman, & Stroessner, 2000). This is because trying to consciously activate or interpret information that is inconsistent to deeply ingrained ideas (i.e., stereotypes) requires the perceiver to carefully attend to the information.

Under this line of reasoning, it is possible to speculate that the interpretation of gender biased role nouns may be subject to differ according to participants' individual cognitive capacity (as investigated in Gygax et al., 2012) as well as experimentally imposed cognitive load manipulations. Studies in the future may want to examine individual cognitive capacities or employ cognitive load tasks to better understand the mechanisms underlying social stereotyping and language.

7.5 Conclusion

The studies presented in this thesis contribute to better understanding how gender information encoded in languages impacts the on-line construction of mental representations. When readers were faced with a gender stereotypical role noun, the encounter of this categorical label automatically activated information pertaining to the role noun's probabilistic gender. In cases where grammatical gender information was available, grammatical information overruled any competing activated conceptual information. Specific gender biases may arise in this case if surface forms of the grammatical gender add to the association of a specific male gender, which in the context of this thesis arose from the ambiguous masculine form that held several interpretation possibilities. Such activation of the specific male gender was robust prevailing over the readers' representations.

These findings on gender information activation were argued as providing evidence for the thinking-for-speaking hypothesis (Slobin, 1996a, 1996b, 2003), which contributed to the notion that when specific aspects of linguistic information are encoded (or omitted) within a language, these characteristics and regularities will become heightened to the reader. Importantly for the current thesis, readers were found to alter their gender biases that conformed to the specific language-bound regularities when reading in that language. In cases where bilinguals switched their representation tendencies to the L2, the magnitude of the representational shift was found to vary according to bilinguals' L2 proficiency. A higher level of L2 proficiency suggested that their representations resembled that of the L1 native readers of the target language, whereas low levels of L2 proficiency indicated a greater reliance of their L1, transferring representation tendencies to their L2. Alternatively, if the linguistic feature in L2 did not exist in bilinguals' L1, the impact of this feature was found to be reduced due to the difficulty in its acquisition.

Future research should pursue investigation of the comprehension process involved in the representation of gender information as they provide a fundamental understanding of the impact of language on cognitive processes.

REFERENCES

- Alderson, J. C. (1984). Reading in a foreign language: A reading problem or a language problem? In J. C. Alderson & A. H. Urquhart (Eds.), *Reading in a foreign language* (pp. 1–24). London: Longman.
- Allport, G. W. (1954). *The nature of prejudice*. Cambridge, M.A.: Addison-Wesley Pub. Co.
- American Psychological Association. (2001). *Publication manual of the American Psychological Association*. Washington, D.C.: American Psychological Association.
- Andersen, S. M., & Klatzky, R. L. (1987). Traits and social stereotypes: Levels of categorization in person perception. *Journal of Personality and Social Psychology*, 53, 235–246. <http://doi.org/10.1037/0022-3514.53.2.235>
- Andersen, S. M., Klatzky, R. L., & Murray, J. (1990). Traits and social stereotypes: Efficiency differences in social information processing. *Journal of Personality and Social Psychology*, 59(2), 192–201. <http://doi.org/10.1037/0022-3514.59.2.192>
- Athanasopoulos, P. (2006). Effects of the grammatical representation of number on cognition in bilinguals. *Bilingualism: Language and Cognition*, 9, 89–96. <http://doi.org/10.1017/S1366728905002397>
- Athanasopoulos, P. (2009). Cognitive representation of colour in bilinguals: The case of Greek blues. *Bilingualism: Language and Cognition*, 12, 83–95. <http://doi.org/10.1017/S136672890800388X>
- Athanasopoulos, P., Damjanovic, L., Krajciova, A., & Sasaki, M. (2011). Representation of colour concepts in bilingual cognition: The case of Japanese blues. *Bilingualism: Language and Cognition*, 14, 9–17. <http://doi.org/10.1017/S1366728909990046>
- Athanasopoulos, P., & Kasai, C. (2008). Language and thought in bilinguals: The case of grammatical number and nonverbal classification preferences. *Applied Psycholinguistics*, 29, 105–123. <http://doi.org/doi:10.1017/S0142716408080053>
- Baayen, R. H. (2013). languageR: Data sets and functions with “Analyzing linguistic data: A practical introduction to statistics” (Version R package version 1.4.1.). Retrieved from <http://CRAN.R-project.org/package=languageR>
- Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59, 390–412. <http://doi.org/10.1016/j.jml.2007.12.005>
- Banaji, M., & Greenwald, A. G. (1995). Implicit gender stereotyping in judgments of fame. *Journal of Personality and Social Psychology*, 68, 181–198. <http://doi.org/10.1037/0022-3514.68.2.181>
- Banaji, M., & Hardin, C. (1996). Automatic stereotyping. *Psychological Science*, 7, 136–141. <http://doi.org/10.1111/j.1467-9280.1996.tb00346.x>

- Barber, H., & Carreiras, M. (2005). Grammatical gender and number agreement in Spanish: an ERP comparison. *Journal of Cognitive Neuroscience*, 17, 137–153.
<http://doi.org/10.1162/0898929052880101>
- Bargh, J. A. (1999). The cognitive monster: The case against the controllability of automatic stereotype effects. In S. Chaiken & Y. Trope (Eds.), *Dual-process theories in social psychology* (pp. 361–382). New York: Guilford Press.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230–244. <http://doi.org/10.1037/0022-3514.71.2.230>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3). <http://doi.org/10.1016/j.jml.2012.11.001>
- Barsalou, L. W. (1993). Flexibility, structure, and linguistic vagary in concepts: Manifestations of a compositional system of perceptual symbols. In A. F. Collins, S. Gathercole, M. A. Conway, & P. E. Morris (Eds.), *Theories of memory* (pp. 29–101). London: Lawrence Erlbaum.
- Bassetti, B. (2007). Bilingualism and thought: Grammatical gender and concepts of objects in Italian-German bilingual children. *International Journal of Bilingualism*, 11, 251–273.
<http://doi.org/10.1177/13670069070110030101>
- Bassetti, B. (2011). The grammatical and conceptual gender of animals in second language users. In V. J. Cook & B. Bassetti (Eds.), *Language and bilingual cognition*, (pp. 357–84). Hove, U.K.: Psychology Press.
- Bates, D. M., Maechler, M., Bolker, B., & Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4* [R package version 1.1-7]. Retrieved from <http://CRAN.R-project.org/package=lme4>.
- Bem, S. L., & Bem, D. J. (1973). Does sex-biased job advertising “aid and abet” sex discrimination? *Journal of Applied Social Psychology*, 3, 6–18.
<http://doi.org/10.1111/j.1559-1816.1973.tb01290.x>
- Berman, R. A., & Slobin, D. I. (1994). *Relating events in narrative: Acrosslinguistic developmental study*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Bernhardt, E. B., & Kamil, M. L. (1995). Interpreting relationships between L1 and L2 reading: Consolidating the linguistic threshold and the linguistic interdependence hypotheses. *Applied Linguistics*, 16, 15–34. <http://doi.org/10.1093/applin/16.1.15>
- Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers’ conceptions of time. *Cognitive Psychology*, 43, 1–22.
<http://doi.org/http://dx.doi.org/10.1006/cogp.2001.0748>
- Boroditsky, L., Schmidt, L., & Phillips, W. (2003). Sex, syntax, and semantics. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and cognition* (pp. 61–80). Cambridge, MA: MIT Press.

- Bower, G. H., & Karlin, M. B. (1974). Depth of processing pictures of faces and recognition memory. *Journal of Experimental Psychology*, 103, 751–757.
<http://doi.org/10.1037/h0037190>
- Brauer, M., & Landry, M. (2008). Un ministre peut-il tomber enceinte? L’impact du générique masculin sur les représentations mentales [Can a secretary of state become pregnant? The impact of the generic masculine on mental representations]. *l’Année psychologique*, 108, 243–272. <http://doi.org/10.4074/S0003503308002030>
- Braun, F., Gottburgsen, A., Sczesny, S., & Stahlberg, D. (1998). Können Geophysiker Frauen sein? Generische Personenbezeichnungen im Deutschen [Can “Geophysiker” (geophysicists) be women? Generic terms for describing persons in German]. *Zeitschrift Für Germanistische Linguistik*, 26, 265–283. <http://doi.org/10.1515/zfgl.1998.26.3.265>
- Braun, F., Sczesny, S., & Stahlberg, D. (2005). Cognitive effects of masculine generics in German: An overview of empirical findings. *Communications*, 30, 1–21.
<http://doi.org/10.1515/comm.2005.30.1.1>
- Brown, A., & Gullberg, M. (2008). Bidirectional crosslinguistic influence in L1-L2 encoding of manner in speech and gesture: A study of Japanese speakers of English. *Studies in Second Language Acquisition*, 30, 225–251. <http://doi.org/10.1017/S0272263108080327>
- Bruner, J. S. (1957). On perceptual readiness. *Psychological Review*, 64(2), 123–152.
<http://doi.org/10.1037/h0043805>
- Brysbaert, Fias, W., & Noël, M. P. (1998). The Whorfian hypothesis and numerical cognition: is “twenty-four” processed in the same way as “four-and-twenty”? *Cognition*, 66, 51–77.
[http://doi.org/10.1016/S0010-0277\(98\)00006-7](http://doi.org/10.1016/S0010-0277(98)00006-7)
- Brysbaert, M. (2007). “The language-as-fixed effect fallacy”: Some simple SPSS solutions to a complex problem. Royal Holloway, University of London.
- Bylund, E., & Jarvis, S. (2011). L2 effects on L1 event conceptualization. *Bilingualism: Language and Cognition*, 14, 47–59. <http://doi.org/10.1017/S1366728910000180>
- Cacciari, C., Corradini, P., Padovani, R., & Carreiras, M. (2011). Pronoun resolution in Italian: The role of grammatical gender and context. *Journal of Cognitive Psychology*, 23, 416–434.
<http://doi.org/10.1080/20445911.2011.526599>
- Cacciari, C., & Padovani, R. (2007). Further evidence of gender stereotype priming in language: Semantic facilitation and inhibition in Italian role nouns. *Applied Psycholinguistics*, 277–293. <http://doi.org/10.1017/S0142716407070142>
- Carreiras, M., & Clifton, C. (2004). *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond*. New York: Psychology Press.
- Carreiras, M., Garnham, A., Oakhill, J., & Cain, K. (1996). The use of stereotypical gender information in constructing a mental model: evidence from English and Spanish. *The Quarterly Journal of Experimental Psychology*, 49A, 639–663.
<http://doi.org/10.1080/027249896392531>

- Carrell, P. L. (1991). Second language reading: Reading ability or language proficiency? *Applied Linguistics*, 12, 159–179.
- Carter, J. D., Hall, J. A., Carney, D. R., & Rosip, J. C. (2006). Individual differences in the acceptance of stereotyping. *Journal of Research in Personality*, 40, 1103–1118. <http://doi.org/10.1016/j.jrp.2005.11.005>
- Chatard, A., Guimont, S., & Martinot, D. (2005). Impact de la féminisation lexicale des professions sur l'auto-efficacité des élèves: Une remise en cause de l'universalisme masculin? [Occupational self-efficacy as a function of grammatical gender in French]. *L'Année psychologique*, 105, 249–272. <http://doi.org/10.3406/psy.2005.29694>
- Clahsen, H., Balkhair, L., Schutter, J.-S., & Cunnings, I. (2013). The time course of morphological processing in a second language. *Second Language Research*, 29(1), 7–31. <http://doi.org/10.1177/0267658312464970>
- Clarke, M. A. (1979). Reading in Spanish and English: Evidence from adult ESL students. *Language Learning*, 29, 121–150. <http://doi.org/10.1111/j.1467-1770.1979.tb01055.x>
- Clark, H. H. (1973). The language-as-fixed-effect fallacy: A critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, 12, 335–359. [http://doi.org/10.1016/S0022-5371\(73\)80014-3](http://doi.org/10.1016/S0022-5371(73)80014-3)
- Cohen, J. D., MacWhinney, B., Flatt, M. R., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, 25, 257–271. <http://doi.org/10.3758/BF03204507>
- Coleman, J. A. (1994). Profiling the advanced language learner: the C-Test in British further and higher education. In R. Grotjahn (Ed.), *Der C-Test. Theoretische Grundlagen und Praktische Anwendungen [The C-test: theoretical foundations and practical applications]*. (Vol. 2, pp. 217–237). Bochum: Brockmeyer.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82, 407–428. <http://doi.org/10.1037/0033-295X.82.6.407>
- Cook, V. J., & Bassetti, B. (2010). *Language and bilingual cognition*. New York and Hove: Psychology Press.
- Corbett, G. G. (1991). *Gender*. Cambridge: Cambridge University Press.
- Costa, A., Caramazza, A., & Sebastian-Galles, N. (2000). The cognate facilitation effect: implications for models of lexical access. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 26(5), 1283–1296.
- Council of Europe. (n.d.). Retrieved from http://www.coe.int/t/dg4/linguistic/cadre1_en.asp
- Curt Hoffman, Lau, I., & Johnson, D. R. (1956). The linguistic relativity of person cognition: An English-Chinese comparison. *Journal of Personality and Social Psychology*, 51, 1097–1105.
- Dahan, D., Swingle, D., Tanenhaus, M. K., & Magnuson, J. S. (2000). Linguistic gender and spoken-word recognition in French. *Journal of Memory and Language*, 42, 465–480. <http://doi.org/10.1006/jmla.1999.2688>

- Daller, H., & Phelan, D. (2006). The C-test and TOEIC as measures of students' progress in intensive short courses in EFL. In R. Grotjahn (Ed.), *Der C-Test: Theorie, Empirie, Anwendungen. The C-Test: Theory, empirical research, applications. Volume 6.* (pp. 101–119). Peter Lang.
- Danziger, S., & Ward, R. (2010). Language changes implicit associations between ethnic groups and evaluation in bilinguals. *Psychological Science*, *21*, 799–800. <http://doi.org/10.1177/0956797610371344>
- de Groot, A. M. B., Delmaar, P., & Lupker, S. J. (2000). The processing of interlexical homographs in translation recognition and lexical decision: Support for non-selective access to bilingual memory. *The Quarterly Journal of Experimental Psychology*, *53A*, 397–428. <http://doi.org/10.1080/027249800390547>
- Desrochers, A. (1986). Genre grammatical et classification nominale. [Grammatical gender and noun classification.]. *Canadian Journal of Psychology / Revue Canadienne de Psychologie*, *40*, 224–250. <http://doi.org/10.1037/h0080095>
- Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*, 5–18. <http://doi.org/10.1037/0022-3514.56.1.5>
- Dewaele, J.-M., & Véronique, D. (2001). Gender assignment and gender agreement in advanced French interlanguage: a cross-sectional study. *Bilingualism: Language and Cognition*, *4*, 275–297. <http://doi.org/10.1017/S136672890100044X>
- Diekman, A. B., & Eagly, A. H. (2000). Stereotypes as dynamic constructs: Women and men of the past, present, and future. *Personality and Social Psychology Bulletin*, *26*, 1171–1188. <http://doi.org/10.1177/0146167200262001>
- Dijkstra, T., Grainger, J., & van Heuven, W. J. B. (1999). Recognition of cognates and interlingual homographs: The neglected role of phonology. *Journal of Memory and Language*, *41*, 496–518. <http://doi.org/10.1006/jmla.1999.2654>
- Dijkstra, T., & van Heuven, W. J. B. (1998). The BIA model and bilingual word recognition. In J. Grainger & A. M. Jacobs (Eds.), *Localist connectionist approaches to human cognition* (pp. 189–225). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
- Dijkstra, T., Van Jaarsveld, H., & Brinke, S. T. (1998). Interlingual homograph recognition: effects of task demands and Language intermixing. *Bilingualism: Language and Cognition*, *1*, 51–66. <http://doi.org/10.1017/S1366728998000121>
- Dovidio, J. F., Evans, N., & Tyler, R. B. (1986). Racial stereotypes: The contents of their cognitive representations. *Journal of Experimental Social Psychology*, *22*, 22–37. [http://doi.org/10.1016/0022-1031\(86\)90039-9](http://doi.org/10.1016/0022-1031(86)90039-9)
- Duden. (2009). *Die Grammatik* (8th ed., Vol. 4). Mannheim & Leipzig & Wien & Zürich: Dudenverlag.

- Duffy, S., & Keir, J. (2004). Violating stereotypes: Eye movements and comprehension processes when text conflicts with world knowledge. *Memory & Cognition*, 32, 551–559. <http://doi.org/10.3758/BF03195846>
- Eagly, A. H. (1987). *Sex differences in social behavior: A social-role interpretation*. Hillsdale, N.J: Lawrence Erlbaum.
- Eagly, A. H., & Wood, W. (1999). The origins of sex differences in human behavior: Evolved dispositions versus social roles. *American Psychologist*, 54, 408–423. <http://doi.org/10.1037/0003-066X.54.6.408>
- Eckes, T., & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing*, 23(3), 290–325. <http://doi.org/10.1191/0265532206lt330oa>
- Esaulova, Y., Reali, C., & von Stockhausen, L. (2013). Influences of grammatical and stereotypical gender during reading: eye movements in pronominal and noun phrase anaphor resolution. *Language and Cognitive Processes*, 29, 781–803. <http://doi.org/10.1080/01690965.2013.794295>
- Fausey, C. M., Long, B. L., Inamori, A., & Boroditsky, L. (2010). Constructing agency: The role of language. *Frontiers in Psychology*, 1. <http://doi.org/10.3389/fpsyg.2010.00162>
- Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., Vol. 2, pp. 357–411). New York: McGraw-Hill.
- Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.
- Flaherty, M. (2001). How a language gender system creeps into perception. *Journal of Cross-Cultural Psychology*, 32, 18–31. <http://doi.org/10.1177/0022022101032001005>
- Franceschina, F. (2001). Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research*, 17, 213 –247. <http://doi.org/10.1177/026765830101700301>
- Franklin, A., Clifford, A., Williamson, E., & Davies, I. (2005). Color term knowledge does not affect categorical perception of color in toddlers. *Journal of Experimental Child Psychology*, 90(2), 114–141. <http://doi.org/10.1016/j.jecp.2004.10.001>
- Gabriel, U. (2008). Language policies and in-group favoritism: The malleability of the interpretation of generically intended masculine forms. *Social Psychology*, 39, 103–107. <http://doi.org/10.1027/1864-9335.39.2.103>
- Gabriel, U., & Gygax, P. (2008). Can societal language amendments change gender representation? The case of Norway. *Scandinavian Journal of Psychology*, 49, 451–457. <http://doi.org/10.1111/j.1467-9450.2008.00650.x>
- Gabriel, U., Gygax, P., Sarasin, O., Garnham, A., & Oakhill, J. (2008). Au pairs are rarely male: Norms on the gender perception of role names across English, French, and German. *Behavior Research Methods*, 40, 206–212. <http://doi.org/10.3758/BRM.40.1.206>
- Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Hove: Psychology Press.

- Garnham, A., Gabriel, U., Sarrasin, O., Gygax, P., & Oakhill, J. (2012). Gender representation in different languages and grammatical marking on pronouns: when beauticians, musicians, and mechanics remain men. *Discourse Processes*, 49, 481–500.
<http://doi.org/10.1080/0163853X.2012.688184>
- Garnham, A., & Oakhill, J. (1996). The mental models theory of language comprehension. In B. K. Britton & A. C. Graesser (Eds.), *Models of understanding text* (pp. 313–339). Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Garnham, A., Oakhill, J., & Cain, K. (1997). The interpretation of anaphoric noun phrases time course, and effects of overspecificity. *The Quarterly Journal of Experimental Psychology Section A*, 50, 149–162. <http://doi.org/10.1080/713755687>
- Garnham, A., Oakhill, J., & Reynolds, D. (2002). Are inferences from stereotyped role name to characters' gender made elaboratively? *Memory & Cognition*, 30, 439–446.
<http://doi.org/10.3758/BF03194944>
- Gennari, S. P., Sloman, S. A., Malt, B. C., & Fitch, W. T. (2002). Motion events in language and cognition. *Cognition*, 83, 49–79. [http://doi.org/10.1016/S0010-0277\(01\)00166-4](http://doi.org/10.1016/S0010-0277(01)00166-4)
- Gleitman, L. R., & Papafragou, A. (2013). Relations between language and thought. In D. Reisberg (Ed.), *Handbook of cognitive psychology* (pp. 504–523). New York: Oxford University Press.
- Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, 101, 371–395. <http://doi.org/10.1037//0033-295X.101.3.371>
- Grosjean, F., Dommergues, J. Y., Cornu, E., Guillelmon, D., & Besson, C. (1994). The gender-marking effect in spoken word recognition. *Perception & Psychophysics*, 56(5), 590–598.
<http://doi.org/10.3758/BF03206954>
- Grotjahn, R., Klein-Braley, C., & Raatz, U. (2002). C-Tests: An overview. In J. A. Coleman, R. Grotjahn, & U. Raatz (Eds.), *University language testing and the C-Test* (pp. 93–114). Bochum: AKS-Verlag.
- Gygax, P., & Gabriel, U. (2008). Can a group of musicians be composed of women? Generic interpretation of French masculine role names in the absence and presence of feminine forms. *Swiss Journal of Psychology*, 67, 143–151.
<http://doi.org/10.1024/1421-0185.67.3.143>
- Gygax, P., Gabriel, U., Lévy, A., Pool, E., Grivel, M., & Pedrazzini, E. (2012). The masculine form and its competing interpretations in French: When linking grammatically masculine role names to female referents is difficult. *Journal of Cognitive Psychology*, 24, 395–408.
<http://doi.org/10.1080/20445911.2011.642858>
- Gygax, P., Gabriel, U., Sarrasin, O., Oakhill, J., & Garnham, A. (2008). Generically intended, but specifically interpreted: When beauticians, musicians, and mechanics are all men. *Language and Cognitive Processes*, 23, 464–485.
<http://doi.org/10.1080/01690960701702035>

- Hamilton, D. L., & Sherman, J. W. (1994). Stereotypes. In R. S. Wyer & T. K. Srull (Eds.), *Handbook of social cognition* (2nd ed., Vol. 2, pp. 1–68). Hillsdale, N.J.: Lawrence Erlbaum Associates, Inc.
- Hamilton, M. C. (1991). Masculine bias in the attribution of personhood. *Psychology of Women Quarterly*, 15, 393–402. <http://doi.org/10.1111/j.1471-6402.1991.tb00415.x>
- Haviland, S. E., & Clark, H. H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, 13, 512–521. [http://doi.org/10.1016/S0022-5371\(74\)80003-4](http://doi.org/10.1016/S0022-5371(74)80003-4)
- Hess, D. J., Foss, D. J., & Carroll, P. (1995). Effects of global and local context on lexical processing during language comprehension. *Journal of Experimental Psychology: General*, 124, 62–82. <http://doi.org/10.1037/0096-3445.124.1.62>
- Hilton, J. L., & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology*, 47(1), 237–271. <http://doi.org/10.1146/annurev.psych.47.1.237>
- Horiba, Y. (2000). Reader control in reading: Effects of language competence, text type, and task. *Discourse Processes*, 29, 223–267. http://doi.org/10.1207/S15326950dp2903_3
- Hosoda, M., & Stone, D. L. (2000). Current gender stereotypes and their evaluative content. *Perceptual and Motor Skills*, 90, 1283–1294. <http://doi.org/10.2466/PMS.90.3.1283-1294>
- Hothorn, T., Bretz, F., & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, 50(3), 346–363. <http://doi.org/10.1002/bimj.200810425>
- Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition*, 15(02), 422–433. <http://doi.org/10.1017/S1366728911000678>
- Irmen, L. (2007). What's in a (role) name? Formal and conceptual aspects of comprehending personal nouns. *Journal of Psycholinguistic Research*, 36, 431–456. <http://doi.org/10.1007/s10936-007-9053-z>
- Irmen, L., Holt, D. V., & Weisbrod, M. (2010). Effects of role typicality on processing person information in German: Evidence from an ERP study. *Brain Research*, 1353, 133–144. <http://doi.org/10.1016/j.brainres.2010.07.018>
- Irmen, L., & Knoll, J. (1999). On the use of the grammatical gender of anaphoric pronouns in German. A comparison between Finns and Germans. *Sprache & Kognition*, 18, 123–135. <http://doi.org/10.1024//0253-4533.18.34.123>
- Irmen, L., & Köhncke, A. (1996). Zur Psychologie des “generischen” Maskulinums [On the psychology of the “generic” masculine]. *Sprache & Kognition*, 15, 152–166. <http://doi.org/10.1515/zfgl.33.2-3.212>
- Irmen, L., & Kurovskaja, J. (2010). On the semantic content of grammatical gender and its impact on the representation of human referents. *Experimental Psychology (formerly “Zeitschrift Für Experimentelle Psychologie”)*, 57(5), 367–375. <http://doi.org/10.1027/1618-3169/a000044>

- Irmen, L., & Roßberg, N. (2004). Gender markedness of language: The impact of grammatical and nonlinguistic information on the mental representation of person information. *Journal of Language and Social Psychology*, 23, 272–307. <http://doi.org/10.1177/0261927X04266810>
- Irmen, L., & Schumann, E. (2011). Processing grammatical gender of role nouns: Further evidence from eye movements. *Journal of Cognitive Psychology*, 23(8), 998–1014. <http://doi.org/10.1080/20445911.2011.596824>
- Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language and cognition*. New York: Routledge.
- Johnson-Laird, P. (1983). *Mental models: Towards a cognitive science of language, inferences and consciousness*. Cambridge: Cambridge University Press.
- Kaplan, E., Goodglass, H., Weintraub, S., & Goodglass, H. (1983). *Boston naming test*. Philadelphia: Lea & Febiger.
- Karmiloff-Smith, A. (1979). *A functional approach to child language: A study of determiners and reference*. New York: Cambridge University Press.
- Kennison, S. M., & Trofe, J. L. (2003). Comprehending pronouns: A role for word-specific gender stereotype information. *Journal of Psycholinguistic Research*, 32(3), 355–378. <http://doi.org/10.1023/A:1023543602202>
- Kim, K. H. S., Relkin, N. R., Lee, K., & Hirsch, J. (1997). Distinct cortical areas associated with native and second languages. *Nature*, 388, 171–174. <http://doi.org/10.1038/42218>
- Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363–394. <http://doi.org/10.1037/0033-295X.85.5.363>
- Klein-Braley, C., & Raatz, U. (1984). A survey of research on the C-Test1. *Language Testing*, 1, 134–146. <http://doi.org/10.1177/026553228400100202>
- Koeser, S., Kuhn, E. A., & Sczesny, S. (2014). Just reading? How gender-fair language triggers readers' use of gender-fair forms. *Journal of Language and Social Psychology*, 0261927X14561119. <http://doi.org/10.1177/0261927X14561119>
- Konishi, T. (1993). The semantics of grammatical gender: A cross-cultural study. *Journal of Psycholinguistic Research*, 22, 519–534. <http://doi.org/10.1007/BF01068252>
- Kousta, S., Vinson, D., & Vigliocco, G. (2008). Investigating linguistic relativity through bilingualism: The case of grammatical gender. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 34, 843–858. <http://doi.org/10.1037/0278-7393.34.4.843>
- Kreiner, H., Sturt, P., & Garrod, S. (2008). Processing definitional and stereotypical gender in reference resolution: Evidence from eye-movements. *Journal of Memory and Language*, 58, 239–261. <http://doi.org/10.1016/j.jml.2007.09.003>
- Kroll, J. F., Michael, E., Tokowicz, N., & Dufour, R. (2002). The development of lexical fluency in a second language. *Second Language Research*, 18, 137–171. <http://doi.org/10.1191/0267658302sr201oa>

- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language*, 33(2), 149–174. <http://doi.org/10.1006/jmla.1994.1008>
- Kroll, J. F., & Tokowicz, N. (2005). Models of bilingual representation and processing. *Handbook of bilingualism: Psycholinguistic approaches*, 531–553.
- Kuhn, E. A., & Gabriel, U. (2013). Actual and potential gender-fair language Use: The role of language competence and the motivation to use accurate language. *Journal of Language and Social Psychology*, 0261927X13504297. <http://doi.org/10.1177/0261927X13504297>
- Kurinski, E., & Sera, M. D. (2010). Does learning Spanish grammatical gender change English-Speaking adults' categorization of inanimate objects? *Bilingualism: Language and Cognition*, 14, 203–220. <http://doi.org/10.1017/S1366728910000179>
- Kurtzman, H. S., & MacDonald, M. C. (1993). Resolution of quantifier scope ambiguities. *Cognition*, 48, 243–279. [http://doi.org/10.1016/0010-0277\(93\)90042-T](http://doi.org/10.1016/0010-0277(93)90042-T)
- Lehtonen, M., & Laine, M. (2003). How word frequency affects morphological processing in monolinguals and bilinguals. *Bilingualism: Language and Cognition*, 6, 213–225. <http://doi.org/10.1017/S1366728903001147>
- Lepore, L., & Brown, R. (1997). Category and stereotype activation: Is prejudice inevitable? *Journal of Personality and Social Psychology*, 72, 275–287. <http://doi.org/10.1037/0022-3514.72.2.275>
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.
- Lévy, A., Gygax, P., & Gabriel, U. (2014). Fostering the generic interpretation of grammatically masculine forms: When my aunt could be one of the mechanics. *Journal of Cognitive Psychology*, 26, 27–38. <http://doi.org/10.1080/20445911.2013.861467>
- Liben, L. S., Bigler, R. S., & Krogh, H. R. (2002). Language at work: Children's gendered interpretations of occupational titles. *Child Development*, 73, 810–828. <http://doi.org/10.1111/1467-8624.00440>
- Lucy, J. A., & Gaskins, S. (2003). Interaction of language type and referent type in the development of nonverbal classification preferences. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 465–492). Cambridge, M.A.: MIT Press.
- Lyster, R. (2007). *Learning and teaching languages through content: A counterbalanced approach*. Philadelphia: John Benjamins Publishing Company.
- Maas, A., & Arcuri, L. (1996). Language and stereotyping. In C. N. Macrae & C. Stangor (Eds.), *The foundations of stereotypes and stereotyping* (pp. 193–226). New York: Guilford Press.
- Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology*, 51, 93–120. <http://doi.org/10.1146/annurev.psych.51.1.93>

- Macrae, C. N., Bodenhausen, G. V., Milne, A. B., & Ford, R. L. (1997). On the regulation of recollection: The intentional forgetting of stereotypical memories. *Journal of Personality and Social Psychology*, 72, 709–719. <http://doi.org/10.1037/0022-3514.72.4.709>
- Macrae, C. N., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices : A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology*, 66(1), 37–47. <http://doi.org/10.1037/0022-3514.66.1.37>
- MacWhinney, B. (1997). Second language acquisition and the competition model. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 113–142). Mahwah, NJ: Lawrence Erlbaum Associates.
- Malt, B. C., & Sloman, S. A. (2003). Linguistic diversity and object naming by non-native speakers of English. *Bilingualism: Language and Cognition*, 6, 47–67. <http://doi.org/10.1017/S1366728903001020>
- Misersky, J., Gygax, P. M., Canal, P., Gabriel, U., Garnham, A., Braun, F., ... Sczesny, S. (2013). Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behavior Research Methods*, 1–31. <http://doi.org/10.3758/s13428-013-0409-z>
- New, B., Pallier, C., Brysbaert, M., Ferr, L., Holloway, R., Service, U., & Joliot, H. F. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, 36, 516–524. <http://doi.org/10.3758/BF03195598>
- Nissen, U. K. (2013). Is Spanish becoming more gender fair? A historical perspective on the interpretation of gender-specific and gender-neutral expressions. *Linguistik Online*, 58, 99–116.
- Oakhill, J., Garnham, A., & Reynolds, D. (2005). Immediate activation of stereotypical gender information. *Memory & Cognition*, 33, 972–983. <http://doi.org/10.3758/BF03193206>
- Ogunnaike, O., Dunham, Y., & Banaji, M. R. (2010). The language of implicit preferences. *Journal of Experimental Social Psychology*, 46, 999–1003. <http://doi.org/10.1016/j.jesp.2010.07.006>
- Papafragou, A., Hulbert, J., & Trueswell, J. (2008). Does language guide event perception? Evidence from eye movements. *Cognition*, 108, 155–184. <http://doi.org/10.1016/j.cognition.2008.02.007>
- Papafragou, A., Massey, C., & Gleitman, L. (2002). Shake, rattle, “n” roll: the representation of motion in language and cognition. *Cognition*, 84, 189–219. [http://doi.org/10.1016/S0010-0277\(02\)00046-X](http://doi.org/10.1016/S0010-0277(02)00046-X)
- Paradis, M. (1997). The cognitive neuropsychology of bilingualism. In A. M. B. de Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 331–354). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- Pavlenko, A. (1999). New approaches to concepts in bilingual memory. *Bilingualism: Language and Cognition*, 3, 209–230. <http://doi.org/10.1017/S1366728999000322>

- Pavlenko, A. (2011). *Thinking and speaking in two languages*. Bristol, U.K.: Multilingual Matters.
- Peyer, A., & Wyss, E. L. (1998). "JazzmusikerInnen – Weder Asketen noch Müsli-Fifis" – Feministische Sprachkritik in der Schweiz, ein Überblick ["Jazz-musicians - neither ascetics nor cereals-picker" - Feministic language critic in Switzerland, a review]. *Germanistische Linguistik*, 139/140, 117–154.
- Phillips, W., & Boroditsky, L. (2003). Can quirks of grammatical gender affect the way you think? Grammatical gender and object concepts. In R. Alterman & D. Kirsch (Eds.), *Proceedings of the 25th annual conference of the Cognitive Science Society*. Boston: Cognitive Science Society.
- Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an amazonian indigene group. *Science*, 306, 499–503. <http://doi.org/10.1126/science.1102085>
- Pickering, M. J., Frisson, S., McElree, B., & Traxler, M. (2004). Eye movements and semantic composition. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension: ERP, eye-tracking and beyond* (pp. 33–50). Hove: Psychology Press.
- Pinker, S. (1989). *Learnability and cognition: The acquisition of argument structure*. Cambridge: The MIT Press.
- Prewitt-Freilino, J. L., Caswell, T. A., & Laakso, E. K. (2012). The gendering of language: A comparison of gender equality in countries with gendered, natural gender, and genderless languages. *Sex Roles*, 66, 268–281. <http://doi.org/10.1007/s11199-011-0083-5>
- Prior, A., & MacWhinney, B. (2010). A bilingual advantage in task switching. *Bilingualism: Language and Cognition*, 13, 253–262. <http://doi.org/10.1017/S1366728909990526>
- Quick Placement Test. (2001). Oxford: Oxford University Press.
- Pyykkönen, P., Hyönä, J., & van Gompel, R. P. G. (2010). Activating gender stereotypes during online spoken language processing: Evidence from visual world eye tracking. *Experimental Psychology*, 57, 126–133. <http://doi.org/10.1027/1618-3169/a000002>
- Raatz, U., & Klein-Braley, C. (1982). The C-test- a modification of the cloze procedure. In T. Culhane, C. Klein-Braley, & D. K. Stevenson (Eds.), *Practice and problems in language testing. University of Essex occasional paper* (Vol. 4, pp. 113–138). Colchester: University of Essex, Department of Language and Linguistics.
- Rahimi, M., & Saadat, M. (2005). A verbal protocol analysis of a C-test. *Iranian Journal of Applied Linguistics*, 8, 55–85.
- R Core Team. (2013). *R: A language and environment for statistical computing*. [R Foundation for Statistical Computing]. Vienna, Austria. Retrieved from <http://www.R-project.org/>
- Realí, C., Esaulova, Y., & von Stockhausen, L. (2015). Isolating stereotypical gender in a grammatical gender language: Evidence from eye movements. *Applied Psycholinguistics*, 36, 977–1006. <http://doi.org/10.1017/S0142716414000010>

- Reynolds, D., Garnham, A., & Oakhill, J. (2006). Evidence of immediate activation of gender information from a social role name. *Quarterly Journal of Experimental Psychology*, 59, 886–903. <http://doi.org/10.1080/02724980543000088>
- Rothermund, K. (1998). Automatische geschlechtsspezifische Assoziationen beim Lesen von Texten mit geschlechtseindeutigen und generisch maskulinen Text-Subjekten [Automatic gender-specific associations to texts containing gender-specific and masculine generic text subjects]. *Sprache & Kognition*, 17, 183–198.
- Sabourin, L., Stowe, L. A., & de Haan, G. J. (2006). Transfer effects in learning a second language grammatical gender system. *Second Language Research*, 22, 1–29. <http://doi.org/10.1191/0267658306sr259oa>
- Sanford, A. J. (1985). *Cognition and cognitive psychology*. New York: Basic Books, Inc.
- Sapir, E. (1985). *Selected writings of Edward Sapir in language, culture and personality*. Berkley and Los Angeles: University of California Press.
- Sato, S., Gygax, P. M., & Gabriel, U. (2013). Gender inferences: Grammatical features and their impact on the representation of gender in bilinguals. *Bilingualism: Language and Cognition*, 16, 792–807. <http://doi.org/10.1017/S1366728912000739>
- Scheutz, M., & Eberhard, K. M. (2004). Effects of morphosyntactic gender features in bilingual language processing. *Cognitive Science*, 28, 559–588. http://doi.org/10.1207/s15516709cog2804_3
- Szcesny, S., Diekmann, A. B., & Twenge, J. M. (2007). Dynamics of sex-role stereotypes. In Y. Kashima, F. Klaus, & P. Freytag (Eds.), *Stereotype dynamics: Language-based approaches to the formation, maintenance, and transformation of stereotypes* (pp. 137–163). Mahwah, NJ: Lawrence Erlbaum.
- Sera, M. D., Berge, C. A. H., & del Castillo-Pintado, J. (1994). Grammatical and conceptual forces in the attribution of gender by English and Spanish speakers. *Cognitive Development*, 9, 261–292. [http://doi.org/10.1016/0885-2014\(94\)90007-8](http://doi.org/10.1016/0885-2014(94)90007-8)
- Sera, M. D., Elieff, C., Forbes, J., Burch, M. C., Rodríguez, W., & Dubois, D. P. (2002). When language affects cognition and when it does not: an analysis of grammatical gender and classification. *Journal of Experimental Psychology. General*, 131, 377–397. <http://doi.org/10.1037//0096-3445.131.3.377>
- Sherman, J. W. (1996). Development and mental representation of stereotypes. *Journal of Personality and Social Psychology*, 70, 1126–1141. <http://doi.org/10.1037/0022-3514.70.6.1126>
- Silveira, J. (1980). Generic masculine words and thinking. *Women's Studies International Quarterly*, 3, 165–178. [http://doi.org/10.1016/S0148-0685\(80\)92113-2](http://doi.org/10.1016/S0148-0685(80)92113-2)
- Silverberg, S., & Samuel, A. G. (2004). The effect of age of second language acquisition on the representation and processing of second language words. *Journal of Memory and Language*, 51, 381–398. <http://doi.org/10.1016/j.jml.2004.05.003>

- Singmann, H., Bolker, B., & Westfall, J. (2015). *afex: Analysis of Factorial Experiments* [R package version 0.13-145.]. Retrieved from <http://CRAN.R-project.org/package=afex>
- Singular Inversions Inc. (2004). FaceGenModeller (Version Version 3.1.4). Retrieved from <http://www.FaceGen.com>.
- Siyanova-Chanturia, A., Pesciarelli, F., & Cacciari, C. (2012). The electrophysiological underpinnings of processing gender stereotypes in language. *PloS One*, 7, e48712. <http://doi.org/10.1371/journal.pone.0048712>
- Slobin, D. I. (1996a). From “thought and language” to “thinking for speaking.” In J. J. Gumperz & S. C. Levinson (Eds.), *Rethinking linguistic relativity* (pp. 70–96). Cambridge: Cambridge University Press.
- Slobin, D. I. (1996b). Two ways to travel: Verbs of motion in English and Spanish. In M. Shibatani & S. A. Thompson (Eds.), *Grammatical constructions: Their form and meaning* (pp. 195–219). Oxford: Clarendon Press.
- Slobin, D. I. (2000). Verbalized events: a dynamic approach to linguistic relativity and determinism. In S. Niemeier & R. Dirven (Eds.), *Evidence for linguistic relativity* (pp. 107–138). Amsterdam/Philadelphia: John Benjamins.
- Slobin, D. I. (2003). Language and thought online: Cognitive consequences of linguistic relativity. In D. Gentner & S. Goldin-Meadow (Eds.), *Language in mind: Advances in the study of language and thought* (pp. 157–192). Cambridge, MA: MIT Press.
- Srull, T. K., & Wyer Jr., R. S. (1983). The role of control processes and structural constraints in models of memory and social judgment. *Journal of Experimental Social Psychology*, 19, 497–521. [http://doi.org/10.1016/0022-1031\(83\)90013-6](http://doi.org/10.1016/0022-1031(83)90013-6)
- Stahlberg, D., Braun, F., Irmen, L., & Sczesny, S. (2007). Representation of the sexes in language. In K. Fiedler (Ed.), *Social communication* (pp. 163–187). New York: Psychology Press.
- Stahlberg, D., & Sczesny, S. (2001). Effekte des generischen Maskulinums und alternativer Sprachformen auf den gedanklichen Einbezug von Frauen. *Psychologische Rundschau*, 52, 131–140. <http://doi.org/10.1026//0033-3042.52.3.131>
- Stahlberg, D., Sczesny, S., & Braun, F. (2001). Name your favorite musician. *Journal of Language and Social Psychology*, 20, 464–469. <http://doi.org/10.1177/0261927X01020004004>
- Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In M. G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 327–342). Oxford: Oxford University Press.
- Steiger-Loerbroks, V., & von Stockhausen, L. (2014). Mental representations of gender-fair nouns in German legal language: An eye-movement and questionnaire-based study. *Linguistische Berichte*, (237), 57–80.

- Stericker, A. (1981). Does this “he” or “she” business really make a difference? The effect of masculine pronouns as generics on job attitudes. *Sex Roles: A Journal of Research*, 7, 637–641. <http://doi.org/10.1007/BF00291751>
- Stout, J. G., & Dasgupta, N. (2011). When he doesn’t mean you: Gender-exclusive language as ostracism. *Personality and Social Psychology Bulletin*, 37, 757–769. <http://doi.org/10.1177/0146167211406434>
- Swim, J. K., Mallett, R., & Stangor, C. (2004). Understanding subtle sexism: detection and use of sexist language. *Sex Roles*, 51, 117–128. <http://doi.org/10.1023/B:SERS.0000037757.73192.06>
- Swoyer, C. (2011). How does language affect thought? In V. J. Cook & B. Bassetti (Eds.), *Language and Bilingual Cognition* (pp. 23–42). New York and Hove: Psychology Press.
- Tanenhaus, M. K., & Carlson, G. N. (1990). Comprehension of deep and surface verb phrase anaphors. *Language and Cognitive Processes*, 5, 257–280. <http://doi.org/10.1080/01690969008407064>
- Thierry, G., Athanasopoulos, P., Wiggert, A., Dering, B., & Kuipers, J.-R. (2009). Unconscious effects of language-specific terminology on preattentive color perception. *Proceedings of the National Academy of Sciences*, 106, 4567–4570. <http://doi.org/10.1073/pnas.0811155106>
- Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, 33, 285–318. <http://doi.org/10.1006/jmla.1994.1014>
- Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.
- Vervecken, D., & Hannover, B. (2012). Ambassadors of gender equality? How use of pair forms versus masculines as generics impacts perception of the speaker. *European Journal of Social Psychology*, 42, 754–762. <http://doi.org/10.1002/ejsp.1893>
- Vervecken, D., Hannover, B., & Wolter, I. (2013). Changing (s)expectations: How gender fair job descriptions impact children’s perceptions and interest regarding traditionally male occupations. *Journal of Vocational Behavior*, 82, 208–220. <http://doi.org/10.1016/j.jvb.2013.01.008>
- Vigliocco, G., & Franck, J. (2001). When sex affects syntax: Contextual influences in sentence production. *Journal of Memory and Language*, 45, 368–390. <http://doi.org/10.1006/jmla.2000.2774>
- Vigliocco, G., Vinson, D. P., Paganelli, F., & Dworzynski, K. (2005). Grammatical gender effects on cognition: implications for language learning and language use. *Journal of Experimental Psychology. General*, 134, 501–520. <http://doi.org/10.1037/0096-3445.134.4.501>

- Wasserman, B. D., & Weseley, A. J. (2009). ¿Qué? Quoi? Do languages with grammatical gender promote sexist attitudes? *Sex Roles*, 61, 634–643. <http://doi.org/10.1007/s11199-009-9696-3>
- Weber-Fox, C. M., & Neville, H. J. (1996). Maturational constraints on functional specializations for language processing: ERP and behavioral evidence in bilingual speakers. *Journal of Cognitive Neuroscience*, 8, 231–256. <http://doi.org/10.1162/jocn.1996.8.3.231>
- White, L., Valenzuela, E., Kozłowska-Macgregor, M., & Leung, Y.-K. I. (2004). Gender and number agreement in nonnative Spanish. *Applied Psycholinguistics*, 25, 105–133. <http://doi.org/10.1017/S0142716404001067>
- Whorf, B. L. (1956). *Language, thought and reality: Selected writings of Benjamin Lee Whorf*. (J. B. Carroll, Ed.). Cambridge, MA: MIT Press.
- Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2003). Expecting gender: An event related brain potential study on the role of grammatical gender in comprehending a line drawing within a written sentence in Spanish. *Cortex*, 39, 483–508. [http://doi.org/10.1016/S0010-9452\(08\)70260-0](http://doi.org/10.1016/S0010-9452(08)70260-0)
- Wicha, N. Y. Y., Orozco-Figueroa, A., Reyes, I., Hernandez, A., de Barreto, L. G., & Bates, E. A. (2005). When zebras become painted donkeys: Grammatical gender and semantic priming interact during picture integration in a spoken Spanish sentence. *Language and Cognitive Processes*, 20, 553–587. <http://doi.org/10.1080/01690960444000241>
- Winawer, J., Witthoft, N., Frank, M. C., Wu, L., Wade, A. R., & Boroditsky, L. (2007). Russian blues reveal effects of language on color discrimination. *Proceedings of the National Academy of Sciences*, 104, 7780–7785. <http://doi.org/10.1073/pnas.0701644104>
- Wyer, N. A., Sherman, J. W., & Stroessner, S. J. (2000). The roles of motivation and ability in controlling the consequences of stereotype suppression. *Personality and Social Psychology Bulletin*, 26, 13–25. <http://doi.org/10.1177/0146167200261002>
- Zwaan, R. A., & Brown, C. M. (1996). The influence of language proficiency and comprehension skill on situation-model construction. *Discourse Processes*, 21(3), 289–327.
- Zwaan, R., & Radvansky, G. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, 123, 162–185. <http://doi.org/10.1037//0033-2909.123.2.162>

APPENDICES

Appendix A – L2 C-tests

Appendix B – Language questionnaires

Appendix C – Experimental materials

Appendix A - L2 C-tests

A1 - C-test passages used to evaluate L2 English proficiency in Study I

Reconstruct the following texts by filling in the blank spaces (dashes). The number of spaces represents the number of letters you need to fill in.

The Lion

The lion is called the king of beasts. Lions a _ _ found liv _ _ _ wild i _ the grass _ _ _ _ of Afr _ _ _ . They hu _ _ smaller ani _ _ _ _ and fe _ _ on th _ _ . There a _ _ no wi _ _ lions i _ Europe, b _ _ there a _ _ captive li _ _ _ in Euro _ _ _ _ zoos. T _ _ male li _ _ is a beau _ _ _ _ animal. Ro _ _ _ his he _ _ he h _ _ a ri _ _ of lo _ _ hair cal _ _ _ a ma _ _ . When t _ _ lion i _ young, t _ _ hair o _ his ma _ _ is yel _ _ _ . When h _ is o _ _ , the ha _ _ is some _ _ _ _ black. T _ _ female li _ _ , or lio _ _ _ _ , does n _ _ have a ma _ _ . Lions a _ _ dangerous ani _ _ _ _ . A lion can kill a man.

In case someone faints

People faint when the normal blood supply to the brain is suddenly cut down. This c _ _ happen i _ they a _ _ surprised o _ shocked b _ sudden ne _ _ or b _ something th _ _ see. So _ _ people fa _ _ _ if th _ _ see oth _ _ _ hurt. So _ _ people fa _ _ _ in cro _ _ _ . Others fa _ _ _ if th _ _ are i _ a room th _ _ is h _ _ and stu _ _ _ . If a person fai _ _ _ while stan _ _ _ _ , lay h _ _ down. I _ his fa _ _ is pa _ _ , lift h _ _ feet. I _ he i _ sitting do _ _ when h _ faints, pl _ _ _ his he _ _ between h _ _ knees. Loo _ _ _ any ti _ _ _ clothing th _ _ might ke _ _ him fr _ _ breathing eas _ _ _ . If possible, place a cold, wet cloth on his forehead.

The Black Sea

The Black Sea gets its name from the color of its water. In win_ _ _ its co_ _ _ is ve_ _ dark. Th_ _ is cau_ _ _ by fo_ _ that set_ _ _ low ov_ _ the ar_ _ and c_ _ off sunl_ _ _ . The Bl_ _ _ Sea i_ 748 mi_ _ _ from ea_ _ to we_ _ ; it i_ 374 mi_ _ _ from no_ _ _ to so_ _ _ . Four coun_ _ _ _ : Russia, Rom_ _ _ _ , Bulgaria, a_ _ Turkey, bor_ _ _ the s_ _ . Several la_ _ _ rivers em_ _ _ into i_ . The dee_ _ _ _ part o_ the s_ _ is i_ its so_ _ _ central reg_ _ _ . Many po_ _ _ line t_ _ sea. Gr_ _ _ , lumber a_ _ sugar a_ _ the ma_ _ exports th_ _ pass thr_ _ _ _ these po_ _ _ . Fishing is good in the Black Sea and supports many of the people on its coasts.

Different food cultures of the world

People in different countries may eat the same food but they prepare it very differently. For exa_ _ _ _ , Chinese so_ _ is th_ _ _ and cl_ _ _ , but Ger_ _ _ soup i_ thick a_ _ heavy. So_ _ people li_ _ raw me_ _ , while oth_ _ _ like me_ _ only i_ it i_ well-cooked. Ma_ _ people li_ _ butter fr_ _ _ and fi_ _ , but th_ _ _ are peo_ _ _ in In_ _ _ who li_ _ it mel_ _ _ into a _ oil bef_ _ _ they e_ _ it. Many people in the East like plain boiled rice, but in some countries people like theirs made into a sweet pudding.

A2 - C-test passages used to evaluate L2 French proficiency in Studies I, II, III

Complétez les lettres manquant dans le texte suivant. Les nombres des espaces sont équivalents avec les nombres des lettres que vous devez remplir.

Devenez une star en espagnol !

Dire deux ou trois mots d'espagnol, presque tout le monde sait faire. Mais télép_ _ _ _ _ à Madrid o_ _ _ _ _ parler poli_ _ _ _ _ avec u_ _ _ _ _ Mexicain, vo_ _ _ _ _ qui e_ _ _ _ _ déjà u_ _ _ _ _ peu pl_ _ _ _ _ compliqué. Ma_ _ _ _ _ ne per_ _ _ _ _ pas cour_ _ _ _ _ et oubl_ _ _ _ _ votre timi_ _ _ _ _ : il n'e_ _ _ _ _ jamais tr_ _ _ _ _ tard po_ _ _ _ _ faire d_ _ _ _ _ progrès. L'espa_ _ _ _ _ est tr_ _ _ _ _ proche d_ _ _ _ _ français. C'est donc pour vous une langue relativement facile à apprendre.

La clé de la réussite aux examens

La réussite à un examen, c'est 50% de connaissances et 50% de résistance physique et nerveuse. Pour réu_ _ _ _ _ , il fa_ _ _ _ _ d'abord êt_ _ _ _ _ en fo_ _ _ _ _ . Mais com_ _ _ _ _ faire ? Comment n_ _ _ _ _ pas s_ _ _ _ _ laisser écr_ _ _ _ _ par l_ _ _ _ _ stress e_ _ _ _ _ les fati_ _ _ _ _ de l'an_ _ _ _ _ scolaire ? Fau_ _ _ _ _ - _ _ _ _ se lan_ _ _ _ _ dans l_ _ _ _ _ jogging, comm_ _ _ _ _ une cu_ _ _ _ _ de vita_ _ _ _ _ ou par_ _ _ _ _ se déte_ _ _ _ _ sur u_ _ _ _ _ plage do_ _ _ _ _ de Thaï_ _ _ _ _ ? Du ca_ _ _ _ _ ! Il faut réagir sagement.

Comment améliorer le système scolaire en France?

Les horaires des cours sont trop longs, les programmes trop ambitieux. Bref, c'e_ _ _ _ _ le sys_ _ _ _ _ éducatif qu' _ _ _ _ _ faut cha_ _ _ _ _ . De l'arg_ _ _ _ _ , ça n_ _ _ _ _ suffit p_ _ _ _ _ ! Même quel_ _ _ _ _ de mill_ _ _ _ _ ne vo_ _ _ _ _ pas arra_ _ _ _ _ notre situ_ _ _ _ _ ! C'est u_ _ _ _ _ réforme tot_ _ _ _ _ qu'il no_ _ _ _ _ faut. I_ _ _ _ _ y a tr_ _ _ _ _ d'élèves p_ _ _ _ _ classe, o_ _ _ _ _ n'a p_ _ _ _ _ assez d_ _ _ _ _ profs, comment fa_ _ _ _ _ ? Proposons alors de réduire les cours de 60 à 40 minutes.

Frigg en mer du Nord

A 250 km des côtes norvégiennes, 200 hommes travaillent pour extraire quelques millions de mètres cubes de gaz. Les hom_ _ _ y affrontent u_ _ nature hos_ _ _ : le broui_ _ _ _ , le fr_ _ _ , mais sur_ _ _ le ve_ _ qui pe_ _ atteindre 200 km/h. S_ _ Frigg, les condi_ _ _ _ de tra_ _ _ _ sont diffi_ _ _ _ : douze heu_ _ _ quotidiennes pen_ _ _ _ quinze jo_ _ _ de su_ _ _ avant d_ pouvoir pre_ _ _ _ un co_ _ _ de trois sema_ _ _ _ . L'hélicoptère, seul lien avec la terre ferme, assure le transport des hommes, des marchandises et du courrier.

A3 - C-test passages used to evaluate L2 German proficiency in Studies II, IV

Ergänzen Sie die fehlenden Buchstaben in den folgenden Texten. Die Anzahl Leerstellen entspricht der Anzahl der Buchstaben, die Sie einsetzen müssen.

Fragen zur Berufswahl

Alte Berufe verschwinden, neue kommen hinzu: Bei d_ _ Berufswahl ha_ _ _ Schüler
im_ _ _ wieder Fra_ _ _ oder Prob_ _ _ . Denn e_ _ gibt ei_ _ große Anz_ _ _ sehr
versch_ _ _ _ _ Berufe, u_ _ es i_ _ nicht ein_ _ _ , die rich_ _ _ _ Wahl z_ treffen.
D_ _ berufliche Zuk_ _ _ _ sollte m_ _ rechtzeitig pla_ _ _ . Dabei ka_ _ es sinn_ _ _ _ sein,
sich beim Arbeitsamt beraten zu lassen. Manchmal hilft auch ein Test zu den persönlichen
Berufsinteressen.

Gesunde Ernährung

In mehreren europäischen Ländern haben sich die Ernährungsgewohnheiten der Menschen
in den letzten Jahren verändert. Ein aktu_ _ _ _ _ Bericht zei_ _ , dass d_ _ Europäer
heu_ _ _ _ _ mehr a_ _ Ernährung u_ _ Gesundheit ach_ _ _ . Sie es_ _ _ mehr Ob_ _ und
Gem_ _ _ , mehr Fi_ _ _ und vi_ _ weniger Flei_ _ _ . Obwohl man_ _ _ Speisen im_ _ _ noch
ziem_ _ _ _ süß, fet_ _ _ oder sal_ _ _ sind, i_ _ ein kla_ _ _ Trend zu gesünderer Ernährung
erkennbar. Vermutlich wird sich dieser Trend noch fortsetzen.

Geschichte der Familie

Familien haben ihre eigene Geschichte. Oft ka_ _ man s_ _ bis z_ einem se_ _ frühen Zeitp_ _ _ zurückverfolgen. D_ _ Älteren erzä_ _ _ gerne v_ _ ihrer Kind_ _ _ und Jug_ _ . Alte Fot_ _ und Bri_ _ enthalten wich_ _ _ Informationen üb_ _ die Vergang_ _ _ , alte Werkz_ _ _ und Masch_ _ _ zeigen, w_ _ Vorfahren gel_ _ _ und gearb_ _ _ haben. So kann man auf unterhaltsame Weise viel über die Geschichte der eigenen Familie erfahren.

Geschicktes Zeitmanagement

Zeitmanagement: In den Ohren vieler Studenten klingt dieses Wort eher abschreckend. Dabei kön_ _ _ gerade ei_ _ gute Pla_ _ _ viele ne_ _ Freiräume scha_ _ _ . Mit Hi_ _ _ eines gen_ _ _ Arbeitsplans ka_ _ man näm_ _ _ die fr_ _ _ Zeit oh_ _ schlechtes Gewi_ _ _ genießen, u_ _ der unange_ _ _ Lernstress v_ _ den Prüf_ _ _ wäre vor_ _ . Die Lös_ _ _ ist ga_ _ einfach: Sei_ _ Arbeit systematisch in Zeitblöcke aufteilen. Auch Pausen sollten im Arbeitsplan aufgeführt werden.

A4 - C-test passages used to evaluate L2 English proficiency in Study III

Reconstruct the following texts by filling in the blank spaces (dashes). The number of spaces represents the number of letters you need to fill in.

Geography

The UK is located on a group of islands known as the British Isles, which lie between the Atlantic Ocean and the North Sea, northwest of France. At its widest t_ _ UK i_ 300 mi_ _ _ across a_ _ 600 mi_ _ _ from No_ _ _ to So_ _ _ . It sha_ _ _ a sin_ _ _ land bor_ _ _ with the Irish Repu_ _ _ . Despite i_ _ relatively sm_ _ _ size t_ _ UK boa_ _ _ incredibly var_ _ _ and of_ _ _ very beau_ _ _ _ scenery, fr_ _ the mountains and valleys of the North and West to the rolling landscape of the South, and from downland and heath to fens and marshland.

UK Passport Service

A new passport office that has opened in London will help the UK Passport Service provide a much better service to customers who need a passport urgently. The n_ _ office ru_ _ on a_ appointment-only ba_ _ _ , removing t_ _ need f_ _ a len_ _ _ _ wait bef_ _ _ being se_ _ . The n_ _ building, Globe House repl_ _ _ _ the Petty France off_ _ _ , which af _ _ _ 50 ye_ _ _ of conti_ _ _ _ _ service, h_ _ now clo_ _ _ its do_ _ _ . The London Passport Office h_ _ the capa_ _ _ _ to issue 5000 passports weekly.

Record Employment

Latest employment figures show that there are 28.2 million people in work. Work & Pensions Secretary Alistair Darling said this showed the UK labour market has coped well so far with the current international economic uncertainty. Mr. Darling said: "Employment continued to rise, with this month's figures showing a record 28.2 million people in work. There are 65,000 more people in work than last quarter and 252,000 more than last year. Although both measures of unemployment have risen slightly, they are still significantly lower than they were a year ago." The latest claimant count figures, for the month on Dec 13 2001, show a rise of 3,200 on the previous month. At 963,500 claimants, it remains 70,000 lower than this time last year.

Government consults on plans to modernise animal welfare

Plans to review, modernise and simplify outdated laws on animal welfare have been announced by the Government. Animal welfare groups, local authority representatives, courts, police and industry are to be consulted in which will be a far-reaching review drawing together the environmental and industrial concerns over animal welfare. The Department for the Environment, Food and Rural Affairs (DEFRA) wants to hear views on the existing 11 Acts of Parliament governing the welfare of pets and farm animals.

Appendix B – Language questionnaires

B1 - Self-evaluation questionnaire for native English speakers

Participant #

(!) You may choose several answers!

1. Which language do you normally speak at home?

2. What language were you educated in?

3. At what age did you start learning French?

4. Through what means have you studied this language?

☐ At school

Academic certificate: _____ Number of years: _____

☐ In a French-speaking school

☐ One or both of my parents is / are (a) native speaker(s)

☐ With another member of the family

☐ Through travels (e.g., Erasmus)

☐ Other: _____

5. Have you had other opportunities to improve your competencies in English in another context?

☐ Language exchange program or trips

☐ French-speaking friends

☐ A bilingual internship

☐ Other: _____

6. Please indicate the context in which French currently holds an important function for you. Please specify in what context (e.g., type of work, university domain, etc.).

☐ Work (Type of work: _____)

☐ Internship (affiliation, domain: _____)

☐ With friends

☐ With family

☐ Place of residence (place: _____)

☐ Other: _____

7. Please evaluate your linguistic competency in the French language.

A1 = beginner level → C2 = advanced level (Please refer to the document in annex.

Listening:	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2
Reading:	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2
Conversation:	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2
Oral expression:	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2
Writing:	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2

B2 - Self-evaluation questionnaire for native French speakers

Participant n° :

(!) Il est possible de cocher plusieurs réponses !

1. Quelle est la langue que vous parlez principalement à la maison ?

2. Quelle a été votre langue scolaire principale ?

3. A quel âge avez-vous commencé à apprendre l'allemand ?

4. Par quel moyen avez-vous commencé l'apprentissage de cette langue ?

- ☐ À l'école
Degré scolaire : _____ Nombre d'années : _____
- ☐ Dans une école germanophone
- ☐ Un ou des parents bilingues
- ☐ Un autre membre de la famille
- ☐ Un voyage à l'étranger
- ☐ Autre : _____

5. Par la suite, avez-vous eu l'occasion d'améliorer vos compétences en allemand dans un autre contexte ?

- ☐ Voyage linguistique
- ☐ Ami-e-s germanophones
- ☐ Formation effectuée en bilingue
- ☐ Autre : _____

6. Quel emploi de l'allemand faites-vous actuellement ? (veuillez préciser quel travail, quel domaine universitaire,...)

- ☐ Au travail (emploi : _____)
- ☐ Dans la formation (filière, domaine : _____)
- ☐ Avec des amis
- ☐ Dans la famille
- ☐ Lieu de résidence (lieu : _____)
- ☐ Autre : _____

7. Faites une appréciation de vos compétences linguistiques en allemand.

A1 = connaissance élémentaire → C2 = connaissance avancée (voir feuille annexe pour des précisions)

Ecoute :	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2
Lecture :	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2
Conversation :	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2
Expression orale :	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2
Ecriture :	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2

B3 - Self-evaluation questionnaire for native German speaker

Teilnehmer N° :

(!) Es ist möglich, mehrere Antworten zu wählen!

1. Was ist die Sprache, die Sie am meistens zu Hause sprechen?

2. Was war Ihre Hauptsprache in der Schule?

3. Wie alt waren Sie wenn Sie Französisch lernen angefangen haben?

4. Mit welchen Mitteln haben Sie das Lernen dieser Sprache angefangen?

☐ In die Schule

Stufe: _____ Für wieviel Jahre: _____

☐ In einer französischsprachigen Schule

☐ Ein oder beide Eltern zweisprachig

☐ Ein anderes Familienmitglied

☐ Eine Reise ins Ausland

☐ Andere Mittel : _____

5. Später, haben Sie die Möglichkeit gehabt, Ihre Französisch-Kenntnisse in einem anderen Kontext zu verbessern?

☐ Sprachreise

☐ Frankophonen Freunde

☐ Ausbildung in zweisprachigen durchgeführt

☐ Andere Mittel : _____

6. Was Gebrauch der französischen Sprache haben Sie derzeit? (Bitte angeben: welche Stelle, welche akademischen Bereit,...)

☐ Bei der Arbeit (Stelle : _____)

☐ In der Ausbildung (Studiengang, Gebiet : _____)

☐ Mit Freund

☐ In der Familie

☐ Wohnort (Stadt : _____)

☐ Andere Mittel : _____

7. Machen Sie eine Bestandsaufnahme Ihrer Sprachkenntnisse in Deutsch.
 A1 = Grundkenntnisse → C2 = fortgeschrittene Kenntnisse (beiliegend Blatte für genauere Angabe ansehen)

Hören :	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2
Lesen :	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2
Gespräche :	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2
Mündlicher Ausdruck	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2
Schreiben :	<input type="checkbox"/> A1	<input type="checkbox"/> A2	<input type="checkbox"/> B1	<input type="checkbox"/> B2	<input type="checkbox"/> C1	<input type="checkbox"/> C2

Appendix C – Experimental materials

C1 - Examples of men (on top) and mix (on bottom) pairs of faces used for facial recognition task in Study II



C2 - Experimental passages in English used in Study III

Passages with male stereotype information

1. While working for the secret service,
the spy frequently changed identity.
She / He hoped to never be caught.
2. While trying to get the ball near the hole,
the golfer seemed disappointed.
She / He knew that the trajectory would not be good.
3. While thinking about getting elected,
the politician seemed hopeless.
She / He had thought that the polls would show better results.
4. While apprehending those responsible for the burglary,
the police officer looked pleased.
She / He hoped that the crime rate would decrease.
5. While analysing the mathematical simulation results,
the statistician looked disappointed.
She / He believed that there would be significant results.
6. While specifying the deadline for the financial plan,
the boss looked pleased.
She / He thought that the plan would be successful.
7. While installing the software for the department,
the computer specialist appeared doubtful.
She / He knew that the new version had an unfixable bug.
8. While operating on the patient,
the surgeon appeared concerned.
She / He hoped that the patient would recover
9. While sorting out the computer malfunction,
the technician appeared confident.
She / He believed that the recent bug was the cause.
10. While checking the building blueprints,
the engineer seemed anxious.
She / He hoped that electrical power could be easily set up.

11. While revising the chapter on quantum energy,
the physics student appeared anxious.
She / He knew that the test would not be easy.
12. While entering the cockpit.
the pilot looked worried.
She / He hoped that the skies would be clear that day.

Passages with neutral stereotype information

13. While humming the melody,
the singer seemed pleased.
She / He thought that the performance would entertain the crowd.
14. While striding along the pavement,
the pedestrian seemed happy.
She / He knew that the day would be sunny.
15. While watching the film,
the cinemagoer seemed terrified.
She / He hoped that the scary scene would be over soon.
16. While listening to the song,
the concertgoer seemed satisfied.
She / He knew that the other spectators were also moved by the performance.
17. While finishing some homework,
the schoolchild seemed restless.
She / He knew that she / he had to finish her / his work before being allowed to play.
18. While watching the empty stage,
the spectator looked restless.
She / He hoped that the comedian would be on stage soon.
19. While looking at the plants,
the neighbour appeared fascinated.
She / He hoped that the buds would bloom soon.
20. While paddling at the seashore,
the swimmer was surprised.
She / He believed that the water would be warmer.

21. While taking the trophy,
the tennis player appeared fulfilled.
She / He believed that she / he was on the road to success.
22. While writing the book,
the author appeared unsure.
She / He knew that the readers would not be happy with the ending.
23. While playing a tune on the piano,
the musician appeared happy.
She / He knew that the audience would appreciate the melody
24. While descending the slope,
the skier seemed hesitant.
She / He knew that it would be icier than the day before.

Passages with female stereotype information

25. While giving a facial scrub,
the beautician seemed hesitant.
She / He hoped that it would not be too painful for the client.
26. While wrapping up the new-born baby,
the birth attendant seemed eager.
She / He knew that the mother would be delighted to finally see the baby.
27. While examining the tarot cards,
the fortune teller seemed worried.
She / He hoped that what she / he was seeing was not true.
28. While scanning the grocery bar-codes,
the cashier seemed puzzled.
She / He hoped that she / he would not make any mistakes.
29. While preparing the hospital bed,
the nurse looked optimistic.
She / He hoped that that the patient would get better.

30. While cutting the client's fringe,
the hairdresser looked worried.
She / He knew that she / he had cut too much.
31. While revising the chapter on Freud,
the psychology student looked convinced.
She / He realised that she / he very much liked psychoanalysis.
32. While advising on a low-carb regime,
the dietician looked hesitant.
She / He knew that it was not going to be easy for the client.
33. While shortening the trousers,
the dressmaker seemed confident.
She / He knew that the outfit would fit.
34. While learning the choreography,
the dancer was focused.
She / He knew that this particular sequence would be important.
35. While calculating the bill,
the sales assistant appeared distracted.
She / He looked forward to finishing the day.
36. While proposing better child care services,
the social worker seemed enthusiastic,
She / He knew that there was more she / he could do to help the families.

C3 - Experimental passages in French used in Study III

Passages with male stereotype information

1. En travaillant pour les services secrets,
l'espion changeait souvent d'identité.
Elle / Il espérait ne jamais être découvert.
2. En frappant la balle pour se rapprocher du trou,
le golfeur sembla déçu.
Elle / Il savait que la trajectoire ne serait pas bonne.
3. En s'imaginant être élu,
le politicien semblait désespéré.
Elle / Il croyait que les sondages montreraient de meilleurs résultats.
4. En appréhendant les responsables du cambriolage,
le policier eut l'air content.
Elle / Il espérait que le taux de criminalité diminuerait.
5. En analysant les résultats de la simulation mathématique,
le statisticien avait l'air déçu.
Elle / Il croyait qu'il y aurait des résultats significatifs.
6. En précisant le délai du plan financier,
le patron avait l'air satisfait.
Elle / Il pensa que le plan réussirait.
7. En installant le logiciel pour le département,
l'informaticien paraissait incertain.
Elle / Il savait que la nouvelle version avait un bug irréparable.
8. En opérant le patient,
le chirurgien paraissait préoccupé.
Elle / Il espérait que le patient se remettrait.
9. En réparant la défaillance informatique,
le technicien parut confiant.
Elle / Il croyait que le bug récent en était la cause.
10. En vérifiant les plans de construction,
l'ingénieur sembla anxieux.
Elle / Il espérait que le courant électrique serait facile à établir.

11. En révisant le chapitre sur le quantum d'énergie,
l'étudiant en physique paraissait tendu.
Elle / Il savait que l'examen ne serait pas facile.

12. En entrant dans le cockpit,
l'aviateur eut l'air inquiet.
Elle / Il espérait que le ciel serait clair ce jour-là.

Passages with neutral stereotype information

13. En fredonnant la mélodie,
le chanteur semblait heureux.
Elle / Il pensait que la représentation divertirait le public.

14. En arpentant le trottoir,
le promeneur sembla content.
Elle / Il savait que la journée serait belle.

15. En regardant le film,
le spectateur de cinéma semblait terrifié.
Elle / Il espérait que la scène effrayante se terminerait bientôt.

16. En écoutant la chanson,
l'auditeur de concert semblait satisfait.
Elle / Il savait que les autres spectateurs étaient aussi touchés par la représentation.

17. En finissant ses devoirs,
l'écopier avait l'air impatient.
Elle / Il savait qu'elle / il devait terminer tout son travail avant de pouvoir jouer.

18. En regardant la scène vide,
le spectateur eut l'air agité.
Elle / Il espérait que l'humoriste arriverait bientôt.

19. En regardant les plantes,
le voisin avait l'air fasciné.
Elle / Il espérait que les boutons fleuriraient bientôt.

20. En barbotant dans la mer,
le nageur fut surpris.
Elle / Il croyait que l'eau serait plus chaude.

21. En prenant le trophée,
le joueur de tennis parut satisfait.
Elle / Il pensait qu'il était sur la voie du succès.
22. En écrivant le livre,
l'auteur paraissait incertain.
Elle / Il savait que les lecteurs ne seraient pas contents du dénouement.
23. En jouant un air au piano,
le musicien parut content.
Elle / Il savait que le public apprécierait la mélodie.
24. En descendant la piste,
le skieur paraissait hésitant.
Elle / Il savait qu'elle serait davantage verglacée que la veille.

Passages with female stereotype information

25. En donnant les soins du visage,
l'esthéticien semblait hésitant.
Elle / Il espérait que ce ne serait pas trop douloureux pour la cliente.
26. En enveloppant le nouveau-né,
l'assistant maternel sembla enthousiaste.
Elle / Il savait que la mère serait ravie de voir enfin son bébé.
27. En lisant les tarots,
le diseur de bonne aventure sembla inquiet.
Elle / Il espérait que ce qu'elle / il voyait n'était pas vrai.
28. En saisissant les code-barres des provisions,
le caissier sembla perplexe.
Elle / Il espérait qu'elle / il ne faisait pas d'erreur.
29. En préparant le lit d'hôpital,
l'infirmier avait l'air optimiste.
Elle / Il espérait que le patient se sentirait bien.
30. En coupant la frange du client,
le coiffeur eut l'air inquiet.
Elle / Il savait qu'elle / il avait coupé trop court.

31. En révisant le chapitre sur Freud,
l'étudiant en psychologie avait l'air convaincu.
Elle / Il réalisa qu'elle / il aimait beaucoup la psychanalyse.
32. En conseillant un régime pauvre en sucres,
le diététicien eut l'air hésitant.
Elle / Il savait que ça n'allait pas être facile pour le client.
33. En raccourcissant les pantalons,
le couturier sembla sûr.
Elle / Il savait que la tenue conviendrait.
34. En apprenant la chorégraphie,
le danseur était concentré.
Elle / Il savait que cette séquence particulière serait importante.
35. En comptant la recette du jour,
le vendeur paraissait distrait.
Elle / Il se réjouissait de la fin de journée.
36. En proposant de meilleurs services de garde d'enfants,
l'assistant social semblait enthousiaste.
Elle / Il savait qu'elle / il pouvait faire plus pour aider les familles.